

**The Social Construction of Risk in Trustworthy Digital Repository Audit and Certification**

By

Rebecca D. Frank

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Information)  
in the University of Michigan  
2018

Doctoral Committee:

Professor Elizabeth Yakel, Chair  
Associate Professor Paul Conway  
Professor Paul N. Courant  
Professor Shobita Parthasarathy

Rebecca D. Frank

[frankrd@umich.edu](mailto:frankrd@umich.edu)

ORCID iD: 0000-0003-2064-5140

© Rebecca D. Frank 2018

## **Dedication**

To Graham.

And in memory of Robert Frank and Dean Henricks.

## **Acknowledgments**

I would like to thank my advisor and dissertation committee chair Beth Yakel, who has been an incredible mentor through my doctoral studies and the dissertation process. I am grateful for her guidance, encouragement, and insight. Beth has been generous with her time and support, and has shown me what it means to be both a researcher and a member of an intellectual community. I would also like to thank the members of my dissertation committee, Paul Conway, Paul Courant, and Shobita Parthasarathy for their guidance and feedback on this dissertation.

I would like to thank Megh Marathe and Carl Haynes, who provided valuable assistance in the data analysis process by acting as additional coders for my interview transcripts. I would also like to thank the research participants who contributed to this study. This dissertation would not have been possible without their generous cooperation.

It has been my great fortune to be part of a vibrant research community. I am thankful for the encouragement, inspiration, and unwavering support of many people, including Andrea Barbarin, Matt Burton, Melissa Chalmers, Joanne Evans, Ixchel Faniel, Kristin Fontichiaro, Patricia Garcia, Tamy Guberek, Margaret Hedstrom, Caitlin Holman, Liz Kaziunas, Adam Kriesberg, Allan Martell, Ricky Punzalan, Tonia Sutherland, Kara Suzuka, Andrea Thomer, and David Wallace. I would like to thank Joanne Mihelcic, who has been a wonderful friend, colleague, and writing partner. Her thoughtful feedback, willingness to talk about research

methods, and periodic reminders to think about the big picture have been invaluable throughout the dissertation process.

I would like to thank the staff of the University of Michigan School of Information, who have helped me to navigate the doctoral program. I am particularly grateful to Veronica Falandino and Allison Sweet, and Rebecca Epstein and Barb Smith.

To my parents, Kent and Michelle, and Kathy and Gary, my brother Henry, and my sister Ruthie, I would like to say thank you. Your unconditional support and love have made this possible. To my grandparents, Dorothy and Bob, Lucille and Dean, and Marianne and Larry, thank you for showing me in so many different ways what it means to live an intellectually adventurous life. And finally, to Graham Hukill – thank you is not enough.

This research was funded in part by a Rackham Graduate Student Research Grant.

## Table Of Contents

<b>Dedication .....</b>	<b>ii</b>
<b>Acknowledgments .....</b>	<b>iii</b>
<b>List Of Tables .....</b>	<b>ix</b>
<b>List Of Figures.....</b>	<b>x</b>
<b>List Of Appendices.....</b>	<b>xi</b>
<b>Abstract.....</b>	<b>xii</b>
<b>Chapter 1: Introduction .....</b>	<b>1</b>
<b>1.1 Background and Problem Statement .....</b>	<b>1</b>
<b>1.2 Theoretical Framework .....</b>	<b>11</b>
<b>1.3 Research Questions .....</b>	<b>15</b>
<b>1.4 Research Design Overview .....</b>	<b>15</b>
<b>1.5 Significance .....</b>	<b>16</b>
<b>1.6 Structure.....</b>	<b>17</b>
<b>Chapter 2: Literature .....</b>	<b>19</b>
<b>2.1 Social Construction of Risk .....</b>	<b>21</b>
2.1.1 Definitions: Risk and Risk Perception .....	22
2.1.2 Factors That Influence Risk Perception .....	24
2.1.2.1 Communication: Amplification/Attenuation.....	24

2.1.2.2	Complexity .....	26
2.1.2.3	Expertise .....	27
2.1.2.4	Organizations.....	28
2.1.2.5	Political Culture.....	30
2.1.2.6	Trust.....	32
2.1.2.7	Uncertainty .....	34
2.1.2.8	Vulnerability.....	35
2.1.3	Risk Assessment & Management.....	38
<b>2.2</b>	<b>Digital Preservation and Risk .....</b>	<b>40</b>
2.2.1	Definition: Digital Preservation .....	40
2.2.2	Risk in the Digital Preservation Literature.....	41
<b>2.3</b>	<b>Trusted Digital Repositories.....</b>	<b>44</b>
2.3.1	Understanding Digital Repositories and Trust.....	45
2.3.2	Audit Processes for Trustworthy Digital Repositories.....	47
2.3.2.1	ISO 16363: Trustworthy Repositories Audit and Certification (TRAC) & ISO 16919: Requirements For Bodies Providing Audit And Certification Of Candidate Trustworthy Digital Repositories (PTAB) .....	48
2.3.2.2	DIN 31644: nestor Seal for Trustworthy Digital Archives .....	51
2.3.2.3	Data Seal of Approval .....	52
2.3.2.4	CoreTrustSeal.....	53
2.3.2.5	The European Framework .....	54
<b>2.4</b>	<b>Conclusion.....</b>	<b>55</b>
<b>Chapter 3: Research Methods .....</b>		<b>59</b>
<b>3.1</b>	<b>Pilot Study .....</b>	<b>62</b>
<b>3.2</b>	<b>Analytical Focus .....</b>	<b>63</b>
<b>3.3</b>	<b>Data Collection &amp; Analysis.....</b>	<b>63</b>
3.3.1	Population and Sample.....	65
3.3.1.1	Standard Developers.....	66

3.3.1.2	Auditors .....	67
3.3.1.3	Repository Staff Members from TRAC Certified Repositories .....	68
3.3.2	Interviews .....	78
3.3.2.1	Testing Interview Protocol .....	81
3.3.2.2	Analysis of Interview Data .....	81
3.3.3	Document Analysis .....	82
3.3.3.1	Available Documentation .....	83
<b>3.4</b>	<b>Limitations .....</b>	<b>84</b>
<b>Chapter 4:</b>	<b>Findings .....</b>	<b>87</b>
<b>4.1</b>	<b>Introduction .....</b>	<b>87</b>
<b>4.2</b>	<b>Risk and TRAC .....</b>	<b>89</b>
<b>4.3</b>	<b>Potential Sources of Risk .....</b>	<b>91</b>
4.3.1	Finance .....	91
4.3.1.1	Succession Planning .....	101
4.3.2	Legal .....	109
4.3.3	Organizational Governance .....	117
4.3.4	Repository Processes .....	130
4.3.5	Technical Infrastructure .....	140
<b>4.4</b>	<b>TRAC Audit Process and Factors that Influence the Social Construction of Risk.....</b>	<b>148</b>
4.4.1	Site Visit .....	148
4.4.2	Maintaining Certification .....	153
<b>4.5</b>	<b>Conclusion .....</b>	<b>158</b>
<b>Chapter 5:</b>	<b>Discussion &amp; Conclusion.....</b>	<b>161</b>
<b>5.1</b>	<b>Summary of Findings.....</b>	<b>161</b>
<b>5.2</b>	<b>Risk in Digital Preservation.....</b>	<b>163</b>
<b>5.3</b>	<b>The Social Construction of Risk in TDR Audit &amp; Certification.....</b>	<b>165</b>



5.3.1	Communication .....	165
5.3.2	Expertise .....	166
5.3.3	Uncertainty .....	168
5.3.4	Vulnerability .....	169
5.3.5	Complexity .....	171
5.3.6	Organizations .....	172
5.3.7	Political Culture .....	173
5.3.8	Trust .....	174
5.3.9	Revised Theoretical Model for the Social Construction of Risk in Digital Preservation .....	175
<b>5.4</b>	<b>Theoretical Contributions And Implications For Research, Policy, &amp; Practice .....</b>	<b>176</b>
5.4.1	The Social Construction of Risk in TRAC Audit and Certification .....	177
5.4.2	Repository Management & Trustworthy Digital Repository Certification .....	178
5.4.3	Digital Preservation Policy .....	181
<b>5.5</b>	<b>Future Directions .....</b>	<b>182</b>
<b>Appendices .....</b>		<b>186</b>
<b>References .....</b>		<b>206</b>

## **List Of Tables**

Table 1: Overview Of CRL Audit Report Scores.....	65
Table 2: Overview Of Interviewees.....	66

## **List Of Figures**

Figure 1: Theoretical Model For The Social Construction Of Risk In Digital Preservation.....	13
Figure 2: Risk Management Process (Barateiro, et al., 2010) .....	42
Figure 3: TRAC Certification Process.....	88
Figure 4: Revised Theoretical Model For The Social Construction Of Risk In Digital Preservation.....	176

## **List Of Appendices**

Appendix A: Interview Protocol for Standard Developers .....	187
Appendix B: Interview Protocol for Auditors .....	190
Appendix C: Interview Protocol for Repository Staff .....	193
Appendix D: Interview Protocol: Repository Staff, Audit Manager .....	196
Appendix E: Vignette .....	200
Appendix F: Interview Data Analysis Code Set .....	202

## **Abstract**

This dissertation examines the social construction of risk in trustworthy digital repository (TDR) certification. It focuses on the Trustworthy Repositories Audit and Certification (TRAC) process, which is administered by the Center for Research Libraries and governed by the ISO 16363 standard. This research seeks to understand how standard developers, auditors, and repository staff members construct their understanding of risk, a foundational concept in digital preservation and TDR certification, in the context of a TRAC audit.

In this dissertation, I have developed an analytical framework of risk that draws on eight social factors that influence how people and groups construct their understandings of risk in the context of digital preservation: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. I argue that although digital preservation has been examined as a technical, economic, and organizational phenomenon, it is also social. I also argue that while the digital preservation community has regarded the concept of risk as a discoverable, calculable value, it is in fact socially constructed, and as such research is needed that considers the social context in which the repositories exist and the ways in which social factors may influence how participants understand and behave in response to risk information.

This research employs a mixed methods research design combining in-depth semi-structured interviews with document analysis to examine: (1) how participants in three groups

(i.e., standard developers, auditors, and repository staff members) construct their understanding of risk in the context of a TRAC audit, and (2) to what degree the eight factors from my analytical framework come into play in the audit process.

My findings reveal the TRAC audit process is one in which the actors involved agree on a definition of risk, but differ about whether an audit process based on this definition can determine trustworthiness with regard to long-term digital preservation. My findings demonstrate that while standard developers, auditors, and repository staff generally share an understanding of the major sources of potential risk that face digital repositories, they disagree about whether and how these risks can be mitigated and how mitigation can be proven. Individuals who are more removed from the day-to-day work of the repositories undergoing an audit are more likely to accept well-documented risk identification and mitigation strategies as sufficient evidence of trustworthiness, while repository staff are skeptical that documentation is sufficient evidence of risk assessment and mitigation and thus question whether this will translate to actual trustworthiness for long-term digital preservation.

My findings support the argument that digital preservation should treat risk as a socially constructed phenomenon and consider how social factors contribute to an understanding of risk by participants in the audit and certification of TDRs. I found that communication, expertise, uncertainty, and vulnerability were particularly strong factors that influenced how auditors and repository staff members understood risk in the context of TRAC audit processes.

This research has brought empirical methods to an emerging discipline and has created a set of baseline data about the first wave of TRAC certifications that will lay a foundation for future research.

# Chapter 1: Introduction

## 1.1 Background and Problem Statement

Every year the National Science Foundation spends billions of dollars on research. For example, in 2017 the NSF had a budget of approximately \$7.5 billion and supported the work of 359,000 people (The National Science Foundation, 2018). Since 2011 the NSF has required that all proposals include data management plans including information about plans to deposit data with a repository in order to ensure that it will be disseminated and preserved (National Science Foundation, 2011). Despite these efforts to ensure the longevity of this valuable digital information, it remains fragile:

*“Gone is the promise of preserving knowledge forever. We are replacing books, maps, and audiovisual recordings with computer code that is less stable than human memory itself. Code is rapidly overwritten or rendered obsolete by new code. Digital data are completely dependent on machines to render them accessible to human perception. In turn, those machines are completely dependent on uninterrupted supplies of energy to run the server farms that store and serve digital data.” (Smith Rumsey, 2016, p. 8)*

Yet, access to digital information is a critical underpinning of core values and functions in our society, from open government, to individual rights, to research and scholarship. Digital preservation is about ensuring the viability, sustainability, and accessibility of that digital information over time (Berman, 2008).

Digital preservation involves more than simply avoiding loss. Repositories preserving digital assets need to develop a sustainable organizational structure and financial stability as well as create robust processes to ensure the viability and accessibility of file formats and the long-term storage and management of data. Digital repositories must have the ability to manage risk in all of these areas (C. Anderson, 2005; Garrett & Waters, 1996; Hey, Tansley, & Tolle, 2009). It is important to understand whether the repositories entrusted with valuable digital information are trustworthy because the content that they are responsible for preserving includes valuable and unique governmental data, cultural artifacts, and research data.

Digital preservation consists of a complex set of activities, managed by individuals acting in organizations (S. Hitchcock, Brody, Hey, & Carr, 2007), which take place continuously over time. In 1996, in the *Report of the Task Force on Archiving of Digital Information*, John Garrett and Donald Waters (“Garrett and Waters”) found that long-term preservation of digital information will require the development of new infrastructures to support “trusted organizations capable of storing, migrating and providing access to digital collections” (Garrett & Waters, 1996, p. 40). Garrett and Waters also found that “a process of certification for digital archives is needed to create an overall climate of trust about the prospects of preserving digital information” (Garrett & Waters, 1996, p. 40). This report proposed two possible models for certification of digital repositories: third party audit and certification, and the standards model.

*“There are at least two models of certification. On the one hand, there is the audit model used, for example, to certify official depositories of government documents. The depositories are subject to periodic and rigorous inspection to ensure that they are fulfilling their mission. On the other hand, there is a standards model which operates, for example, in the preservation community. Participants claim to adhere to standards that an appropriate agency has certified as valid and appropriate; consumers then certify by their use whether the products and services actually adhere to the standards. In its call for certified digital archives, the Task Force has not judged the relative merits of applying either of these particular kinds of models of certification.” (Garrett & Waters, 1996, p. 49)*



Of the two models, third party certification has prevailed in the more than 20 years since the Garrett and Waters report. Since that time, the digital preservation community has sought to implement their recommendations, and to develop a structure for certification of trusted digital repositories (TDRs) (e.g., Consultative Committee for Space Data Systems, 2011, 2012; Data Seal of Approval, 2014a; McHugh, Ross, Innocenti, Ruusalepp, & Hoffman, 2008; “nestor Seal for Trustworthy Digital Archives,” 2013). As of 2018, a limited number of repositories have sought to demonstrate their ability to preserve information through one of these emerging certification processes. Those that are certified are deemed “trustworthy” or “trusted.” Implicit in the certification process is the assumption that repositories have demonstrated their ability to manage risk.

One such certification is the Trustworthy Repositories Audit and Certification: Criteria Checklist (TRAC) (Consultative Committee for Space Data Systems, 2012). This certification process is governed by a formal standard, which forms the basis for the audit and certification process (ISO 16363). A second standard (ISO 16919), Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories (PTAB), governs the process by which individuals can become auditors who are authorized to administer the TRAC standard (Consultative Committee for Space Data Systems, 2011). A repository becomes TRAC certified through a formal process in which external auditors assess evidence provided by the repository. Once certified, the repository can call itself a TDR. The TRAC standard forms the principle subject for this dissertation, which asks how stakeholders in the TRAC certification process construct their understanding of risk for digital preservation.

In this dissertation, I focus on TRAC for several reasons. First, as a formal ISO standard TRAC has been developed through a rigorous and transparent process involving input from

experts and community members. The formality of the ISO standard, coupled with the fact that all but one of the audits to date have been conducted by the same organization (i.e., the Center for Research Libraries (CRL)), suggests that the results should be consistent and comparable across CRL TRAC certified repositories. Second, the TRAC audit process is extensive and rigorous. The requirements for this certification are more detailed and strict than other certifications, such as Data Seal of Approval (DSA), which consisted of a 10-point checklist that was reviewed by peers and community members rather than auditors from an external organization. Finally, TRAC is shifting toward a less controlled model. The newly approved PTAB standard and recent training sessions for auditors indicate that CRL is no longer the only organization certifying repositories as trustworthy. Examining the CRL TRAC audit and certification process now will generate baseline data against which future research in this area can be compared. The formality of the standard, coupled with the rigor of the audit process and the timing of this study combine to make TRAC the most appropriate certification process for this research. Further research, as well as alternative audit and certification processes, will be built on TRAC and our understanding of it.

Scholars, including Brian Lavoie and Lorcan Dempsey (2004) have laid out many ways to examine digital preservation, including focusing on digital preservation as a technical, economic, organizational, or social challenge (Lavoie, 2008; Lavoie & Dempsey, 2004). Many in the digital preservation community, however, consider digital preservation to be primarily a technical challenge (e.g., Jantz & Giarlo, 2007). Digital preservation research has focused on specific technical aspects of preservation, such as file formats (Lawrence et al., 2000), system architecture (Barateiro, Antunes, & Borbinha, 2012), the reliability of storage media (Baker et al., 2006), mechanisms for managing distributed backup systems (Maniatis, Roussopoulos, Giuli,

Rosenthal, & Baker, 2005), and the development of technical standards (e.g., Dobratz & Schoger, 2007; Harmsen, 2008). Given this focus on the technology, it is not surprising that much of the literature is composed of practitioner reports. These reports are largely case studies of individual organizations, focusing on how organizations address these technical challenges in hopes of achieving trustworthy status at some point in the future (e.g., Houghton, 2015; Schultz & Gore, 2010).

Within the context of technical threats to digital preservation, a portion of this literature addresses risk. Risk is conceptualized in probabilistic terms and the literature focuses on identifying and categorizing threats or vulnerabilities that particular repositories face, and developing technical solutions to address them (e.g., Barateiro, Antunes, Freitas, & Borbinha, 2010). The existing research in this area, which has focused primarily on technical factors with regard to TDR certification and risk management, highlights important aspects of digital preservation. However, these approaches fail to adequately address social factors that can influence the behavior of the individuals and groups carrying out the work.

There has also been some work addressing the economic challenges in preserving digital information over time. Economic sustainability is a key element of digital preservation (Berman, 2008), and includes “building an economically viable infrastructure, both social and technical, for maintaining valuable data without significant loss or degradation. This includes the whole socio-technical composition of the repository, the short- and long-term value of the material, the costs of undertaking an action, and the recognition that technologies do not sustain digital objects: institutions do, using the available technology” (Bradley, 2007, p. 157). Existing research in this area has focused on the ability of digital repositories to financially support their

work, assessment of financial models, and the long-term viability of their funding sources (e.g., 4C Project, 2016; Berman, 2008; Berman et al., 2010; Bradley, 2007; Lavoie, 2008).

TRAC goes beyond the current funding environment of a repository and also asks that repositories provide, for example, evidence of a succession plan to ensure the viability of their collections beyond the life of the repository itself (Consultative Committee for Space Data Systems, 2012). Existing research about repository certification addresses economic and organizational stability as important elements of risk management for digital repositories, but approaches research in this area at an organizational level rather than considering the roles of individuals within those organizations. Indeed, studies of digital preservation often focus on the repository as a single entity rather than a collection of individual actors (e.g., Dappert & Farquhar, 2009), and much of the literature in this area consists of case studies (e.g., Schultz & Gore, 2010). Approaching digital preservation as an organizational challenge highlights the fact that the work of digital preservation takes place within organizations and that emphasizes the high levels of cooperation and coordination needed to manage a digital repository, but research is needed that considers the social context in which the repositories exist and the ways in which social factors may influence the individuals in a digital repository, whether those individuals are all positioned within one organization, or across member organizations in a consortium.

What we see from these studies on the technical and organizational challenges and economic sustainability is that risk management is a conceptual foundation for digital preservation, an idea with currency for at least the past 20 years (Conway, 1996). These approaches to digital preservation as risk management depend upon the ability of people and organizations to accurately and effectively assess risk. While some of the risks associated with digital preservation activities, such as media deterioration and file format obsolescence (e.g.,

Ohshima, 2010), storage failures (e.g., Vermaaten, Lavoie, & Caplan, 2012), and disasters, attacks, and economic failures (e.g., Barateiro et al., 2010), are known and understood, we know that even when risks are known people are generally poor at judging how to act in response to those risks (Kahneman, 2013). This poor judgment is based in part on the fact that risk is socially constructed and different actors have differing perceptions of risk. Factors that influence these differing perceptions of risk include communication (e.g., Kasperson & Kasperson, 1996), complexity (e.g., Perrow, 1999), expertise (e.g., Tversky & Kahneman, 1974), organizations (e.g., Hutter & Power, 2005a), political culture (e.g., Jasanoff, 1986), trust (e.g., Nelkin, 1989), uncertainty (e.g., van Est, Walhout, & Brom, 2012), and vulnerability (e.g., Olofsson et al., 2014). These varying perceptions lead individuals and organizations to behave differently in response to risks depending on whether their perception amplifies or attenuates those risks

Digital preservation as an academic discipline has engaged with the concept of risk as a knowable, quantifiable figure that technical systems must be designed to overcome (Barateiro et al., 2010). This approach to digital preservation relies on positivistic perspectives and is heavily influenced by computer science (e.g., Barateiro, Antunes, & Borbinha, 2011). Digital preservation scholarship addressing the concept of risk consists mainly of identifying and classifying types of vulnerabilities or threats, and individual case studies describing actions taken by a particular repository (e.g., Vermaaten et al., 2012). Literature in this area, which consists largely of self-produced case studies about specific organizations, treats risk as an objective value and does not engage meaningfully with the notion that perceptions of risk, rather than the risk itself, drive decision-making and action with regard to digital preservation (Ross & McHugh, 2006a). These positivistic attitudes about risk in digital preservation form the basis for an underlying assumption among many digital preservation researchers and practitioners: that the

stakeholders involved in repository audit and certification processes have the same perceptions of risk and therefore interpret the audit documentation in the same way (RLG-OCLC Working Group on Digital Archive Attributes, 2002).

Digital preservation research regarding TDR certification has focused on technical, economic, and organizational factors. This is not sufficient to account for a complex view of the world. As such, a new approach is needed. In this dissertation I have examined TRAC certification as a process that is carried out by individuals within organizations, who are influenced by social factors. Digital preservation challenges, or risks, cannot be considered as merely technical, economic, or organizational. Rather, digital preservation is also a social process in which risks are interpreted by individuals. Their subsequent actions are influenced by social factors that shape their perception. I will focus specifically on the social construction of risk and the factors that influence perceptions of risk among individual stakeholders in the TRAC certification process.

There are several stakeholder groups that are involved in the TRAC certification process: repository staff members, repository leaders (e.g., members of the Board of Directors), leaders of repository parent institutions, auditors, repository users, data depositors, and the developers of the TRAC standard itself. This study will focus on the standard developers, auditors, and repository staff members, as these three groups are most directly involved in the audit process. Standard developers created and maintain the ISO 16363 standard, which governs the audit process. Auditors assess repositories against the criteria in the standard. And repository staff members provide evidence to auditors, primarily in the form of documentation, that their repository's policies and practices meet the criteria described in the standard. The term

stakeholders will be used throughout to refer to these three groups (i.e., standard developers, auditors, and repository staff members) with regard to the TRAC audit and certification process.

If these stakeholders – standard developers, auditors, and repository staff members – do not share the same perceptions of the risks that the repository must manage in order to become a TDR, it is important that the certification process enables those differences to be surfaced and addressed. If not, the outcomes of certification are likely to lack consistency. The goal of TRAC certification is to assess and provide metrics for repositories to gauge their trustworthiness in several areas, enabling comparison of TRAC certified repositories across all dimensions of the standard. The consistency of scoring is critical in order to facilitate this type of comparison.

Indeed, TRAC is just one example of a digital preservation audit process focused on risk management. Data Seal of Approval (DSA) and the nestor Seal for Trustworthy Digital Archives (nestor) also focus on risk management or digital repositories (Data Seal of Approval, 2014a; nestor Certification Working Group, 2013). DSA is an audit and certification process that, like TRAC, asks repositories to respond to a checklist in order to demonstrate trustworthiness through risk management. The certification process is community-based. Members of the DSA community affiliated with existing DSA-approved repositories assess repositories. nestor is an assessment for digital repositories that is based on the DIN 31644 standard (nestor Working Group Trusted Repositories - Certification, 2009). Like TRAC, the nestor assessment process is based on the OAIS model, and audits are conducted by a group of auditors (Consultative Committee for Space Data Systems, 2012a, 2012; Keitel, 2012).

Each of the processes discussed above, TRAC, DSA, and nestor, treat risk as something to be identified, quantified, and managed through structural, financial, or organizational mechanisms. Researchers who examine risk assume that once identified, people will understand

and respond to particular risks in predictable ways (e.g., Fischhoff, 1983; Royal Society (Great Britain) & Study Group on Risk, 1983). This attitude, which is implicit throughout the digital preservation scholarship around risk management (Innocenti, McHugh, & Ross, 2008; McHugh, 2012), largely fails to incorporate well-established theories of risk perception, judgment, and decision-making from other disciplines (e.g., Kahneman, 2013; Slovic, 1987), which argue that even when risks are well-known, social factors influence perception of those risks. This variation in risk perception leads people to behave differently in response to risks. While risk identification is an important step for repositories seeking to demonstrate and improve their ability to preserve information, it is only part of the process. Equally important is the ability to consider how people and groups construct and process risk information and behave in response to that information. This is where this dissertation lies. I assert that digital preservation is a continuous process of managing risk that is driven by social as well as financial, organizational, and technical issues. Understanding the ways in which social factors influence perceptions of risk is key to understanding the ways in which risks are perceived by repository managers and auditors in the TRAC certification process.

In this dissertation I examine the social dimensions of risk. I analyze conceptualizations of risk by repository staff and auditors, and assess how these affect a TRAC audit and, in turn, the consequences for application of the TRAC standard. I examine whether three groups of stakeholders – standard developers, auditors, and repository staff - share the same perceptions of risk, a foundational concept that underlies the TRAC standard.

In studying the social construction of risk, this dissertation does two unique things. First, it changes the focus from the technical, financial, and organizational aspects of risk management to one that incorporates social aspects of risk management as well. And second, it refocuses



digital preservation discussion on the people involved in the process – not to uncover human error per se, but to acknowledge that people underlie all digital preservation processes.

## **1.2 Theoretical Framework**

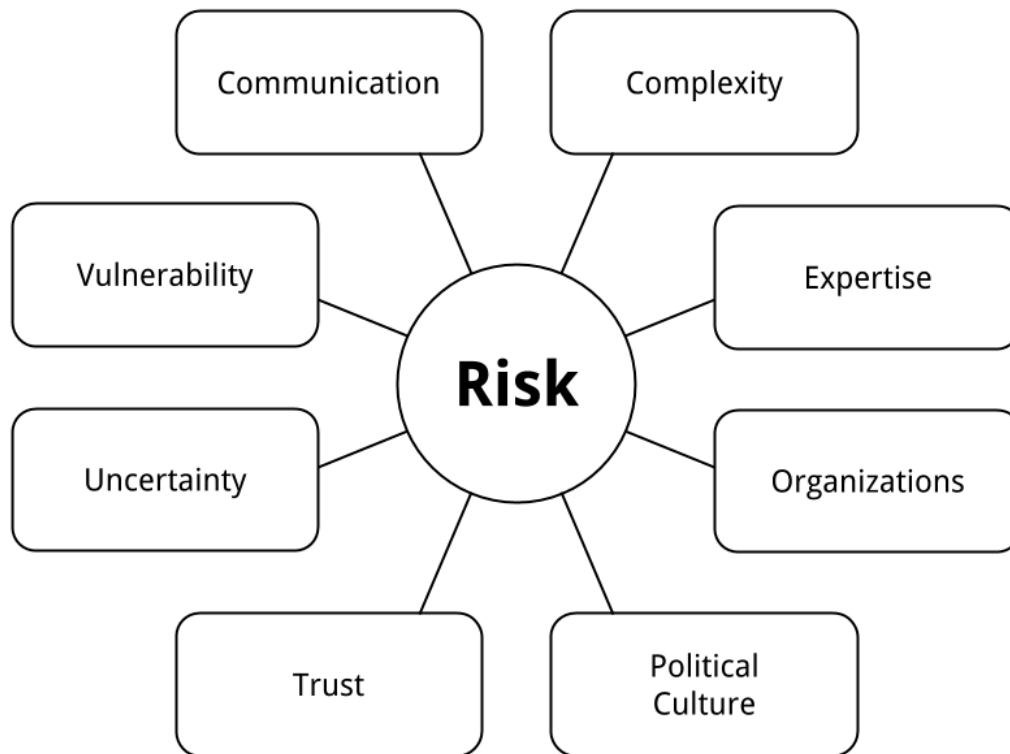
Theories of risk perception form the basis for this study. These theories hold that risk has different meanings for different actors (Renn, 2008), and that social factors influence perceptions of risk (Wilkinson, 2001). Considering risk as a socially constructed concept that influences decisions and actions taken by individuals allows me to consider TDR audit and certification as a socially constructed process (Burgess, 2015). This facilitates an understanding of the audit process as subjective and dependent on the perceptions of the individuals involved.

This framework is crucial to my argument in that people, including repository staff and auditors, whose individual perceptions of risk may influence their actions with regard to audit and certification, conduct the TRAC audit process. One way to understand the process is through the lens of social construction, by considering the social factors that may be influencing the actors involved and how these interact with the organizational, financial, and technical factors (Gergen & Gergen, 2008). Using this lens, I have examined the TRAC certification process in order to critically examine whether the participants in the process (standard developers, auditors, and repository staff) share the same understanding of risk or whether their risk perceptions differ. If perceptions of risk vary across participants (e.g., within a repository, across different repositories, between repository staff and auditors, across different auditors), then the outcomes of the TRAC audits may lack consistency.

A theoretical framework based on factors that influence the social construction of risk would facilitate an examination of the TRAC audit process that is person-centered rather than organizationally-centered and/or materials-centered. Although there are elements of the TRAC

standard that focus on organizational elements (e.g., Section 3: Organizational Infrastructure) or material elements (e.g., Section 5: Infrastructure and Security Risk Management), during the audit process these elements are still addressed by individuals within the organization, who bring their own understandings of risk to bear on the process (Consultative Committee for Space Data Systems, 2012).

In this dissertation, I have developed an analytical framework of risk that draws on eight social factors that influence how people and groups construct their understandings of risk that have been identified in the literature which are relevant for digital repositories: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. The results of this research show that repository staff members disagreed with standard developers and auditors about whether the risk identification and mitigation strategies prescribed in the TRAC standard would translate to actual trustworthiness with regard to long-term digital preservation. These findings have demonstrated that each of the eight factors from the model below contributed to the construction of risk in the TRAC audit and certification process, and that communication, expertise, uncertainty, and vulnerability were the most prominent factors.



**Figure 1: Theoretical Model for the Social Construction of Risk in Digital**

### **Preservation**

Here I describe each factor briefly; they are addressed in greater detail in section 2.1.2 below.

- *Communication*: Perceptions of risk vary depending on the way in which information about those risks is communicated, including the source, method, channel, and means of communication. These elements can either amplify or attenuate perceptions of risk for different individuals and groups (e.g., Bostrom, 2014; Kasperson & Kasperson, 1996).
- *Complexity*: High levels of complexity can make identification difficult with regard to hazards, probabilities, and consequences. Complexity in systems can also lead to unexpected interactions between component parts, often leading to increased levels of risk (e.g., Perrow, 1999; Wilkinson, 2001).

- *Expertise*: Both expertise and lack of expertise can influence perceptions of risk. Experts may have particular knowledge that allows them to understand risk in a particular area, but they have been found to have a narrow focus based on their specialized knowledge, which can influence their perception of risk. Individuals who lack expertise in a particular area may not have the same nuanced understanding of particular areas that experts do, but they have been found to have a greater sense of the broad social context within which they are operating (e.g., E. Vaughan & Seifert, 1992; Wynne, 1992).
- *Organizations*: Organizations both produce and manage risk, and perceptions of risk vary for people depending on their position within an organization. Risk assessment and management activities take place within the context of organizations, and are therefore influenced by the organizations themselves as well as the roles of the individuals within the organizations who participate in those activities (e.g., Hutter, 2005; D. Vaughan, 1996).
- *Political Culture*: National context influences how risks are defined. Perceptions of risk are shaped not only by the political culture within which individuals exist, but also by their place or role within that culture. These factors can elevate or reduce perceptions of risk depending on the position of an individual within the culture. Decisions about how to manage and respond to risks are shaped by political culture as well (e.g., Dake, 1991; Jasanoff, 1986).
- *Trust*: Organizations and processes that involve cooperation by people and groups with different types of knowledge and expertise require trust among those actors. Perceptions of risk can vary depending on the amount of trust that these individuals and groups have for one another (e.g., Nelkin, 1989; Wildavsky & Dake, 1990).
- *Uncertainty*: In many situations it can be difficult to determine and understand risk and its components (hazard, probability, consequences). People and groups operating under conditions of uncertainty may perceive risks differently depending on their level of uncertainty (e.g., Starr, 2003; van Est et al., 2012).
- *Vulnerability*: Risk exposure, or vulnerability, influences perceptions of risk. People and groups who are able to limit their risk exposure may have different perceptions about risk than those who lack the ability to manage their exposure to risks. Greater vulnerability has been shown to increase perceptions of risk, while privilege and the ability to limit or select risk exposure has been shown to decrease perceptions of the severity of risks (e.g., Murphy, 2006; Olofsson et al., 2014).

This study applies theories of risk perception and social construction to the TRAC audit and certification process for digital repositories. These theories suggest that the individuals and groups involved in this process may have different perceptions of risk, a concept that is constructed through social processes (Burgess, 2015), and that these different understandings

influence the consistency of TRAC certification. This is important to consider because TRAC certification is one way that the digital preservation community determines whether repositories are trustworthy with regard to the claims that they make about their ability to preserve digital content long-term.

### **1.3 Research Questions**

My study seeks to answer the following research questions:

1. How do standard developers, auditors, and repository managers conceptualize risk in the context of a TRAC audit?
2. What are the differences and similarities by which standard developers, auditors, and repository managers understand risk as it has been communicated by the TRAC standard?
  - a) In what ways do these differences and similarities become manifest in the TRAC audit process?
3. To what degree do the following eight factors which influence risk perception come into play in the audit process: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability?
  - a) In what ways and why do they emerge when staff and auditors consider risk factors articulated in the TRAC standard?
  - b) What additional factors, if any, emerge which also influence perceptions of risk in relation to the TRAC standard?

### **1.4 Research Design Overview**

In order to address these questions, which were generated from a deep analysis of risk factors in the literature, I conducted a qualitative study to analyze how the concept of risk was constructed in the TRAC audit process. I conducted in-depth semi-structured interviews with three groups of stakeholders: standard developers, auditors, and repository staff members in order to determine whether they share the same perceptions of risk, a foundational concept that underlies the TRAC standard. I critically examined the text of the TRAC standard (ISO 16363),

as well as the audit findings reports published by CRL and other documentation, in order to understand how risk was considered in relation to the standard and applied in the audit process.

## **1.5 Significance**

This research furthers knowledge regarding risk assessment and risk management, and certification processes as applied to digital preservation, data curation, and TDRs. By applying a qualitative design, research findings will (1) bring empirical research to an emerging discipline where scholarship has consisted mainly of case studies produced by practitioners, (2) provide theoretical contributions about how risk perceptions influence assessment of trustworthy digital repositories, (3) develop a theoretical model of factors that influence perceptions of risk with regard to digital preservation, (4) offer insights translatable into the technological, organizational, and economic aspects of any digital repository, thus creating a stronger framework for the long-term preservation and curation of data for reuse, and (5) offer insights translatable into digital preservation policy, including audit and certification processes for trustworthy digital repositories.

This research generates original theoretical contributions to about the social construction of risk in digital preservation. This study expands the ways that we approach digital preservation research by introducing theories of risk perception to the audit and certification process for trusted digital repositories. Existing research in this area has focused on technical definitions of risk that seek to classify types of threats (e.g., Rosenthal, 2010; Vermaaten et al., 2012); an expanded understanding that treats risk as a social construct and includes social and organizational factors is needed. This dissertation extends research about the TRAC audit and certification process by examining the ways in which individual actors may have different perceptions of risk within the process, including standard developers, auditors, and repository

staff. Additionally, existing research in this area has focused on understanding the TRAC audit and certification process at the organizational level (e.g., Kirchhoff, Fenton, Orphan, & Morrissey, 2010; Reilly, Jr. & Waltz, 2013), an understanding of the TRAC process that considers how social factors influence the individual actors involved in the process is needed.

Furthermore, this dissertation has produced and applied a theoretical model for the social construction of risk in digital preservation. While previous research has examined digital preservation as a technical, economic, and organizational phenomenon, this dissertation shows that risk in digital preservation should also be considered as a social phenomenon.

## **1.6 Structure**

This dissertation is organized into five chapters. In Chapter 0 I have outlined the background and summarized the theoretical framework for the study. I have presented my research questions, outlined my research methods, and discussed the significance of this study. In Chapter Chapter 2: I review the literature around three key areas of relevance for this study: the social construction of risk, digital preservation, and trusted digital repository certification. This chapter critically examines existing research in these three areas and identifies a gap in research about trusted digital repositories. Namely, the need for empirical research about TDR certification that considers the ways in which social factors may influence the risk perception of individuals who are involved in the audit and certification process. In Chapter Chapter 3: I describe my research methods, including the analytical focus for the study, methods for data collection and analysis, a description of the population, and a discussion of limitations. In Chapter Chapter 4: I present my findings and argue that although the digital preservation community has regarded the concept of risk as a discoverable value, it is in fact socially constructed. My findings demonstrate that individuals involved in the TRAC audit process view

risk in terms of specific threats, and I explore the five categories of threats that I have identified: finance, legal, organizational governance, repository processes, and technical infrastructure. I also examine two specific aspects of the TRAC audit process in order to understand how the factors from my theoretical model contribute to the social construction of risk. In Chapter Chapter 5:, I discuss how each of the factors from the model for the social construction of risk in digital preservation contributed to the social construction of risk in the TRAC audit process and discuss implications for research, policy, and practice. I conclude by discussing future directions for research.



## **Chapter 2: Literature**

In Chapter Chapter 2:, I introduce relevant literature across three areas: the social construction of risk; digital preservation and risk; and trustworthy digital repository audit and certification processes. In section 2.1, I will contrast the classical risk approach with the social construction of risk. This section describes the differences between a classical definition that characterizes risk as a numeric value determined by considering the probability and magnitude of an adverse event, and a definition of risk which argues that people interpret risks differently as a result of social factors that influence their perceptions. I then discuss the research identifying eight factors that influence perceptions of risk that are particularly relevant for trusted digital repository audit and certification processes, such as TRAC. These are: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. The social construction of risk forms the foundation of the theoretical framework for this study. I have chosen this focus rather than a classical definition of risk because even when risks are known and well-understood, people are poor at judging how to act in response to those risks, and it is the actions of the individual stakeholders in the TRAC audit and certification process that determine the outcomes (Kahneman, 2013; Tversky & Kahneman, 1974).

Section 2.2 examines how the concept of risk has been addressed in the digital preservation literature. While risk is commonly described as being foundational to the field of

digital preservation (e.g., Conway, 1996), it has largely been understood through the classical definition as described in section 2.1.1. Literature in this area has been overwhelmingly practitioner-based, and heavily features practitioner reports rather than empirical research, and has focused on identifying, describing, and classifying particular risks or threats in order to create solutions to manage them. Section 2.3 addresses the concept of trusted digital repositories in the digital preservation literature, beginning with the foundations of TDR certifications in the Garrett & Waters report (Garrett & Waters, 1996). In this section I describe four models for TDR certification: TRAC, nestor, Data Seal of Approval (DSA), and CoreTrustSeal (CTS). As with risk in digital preservation, literature in this area heavily features practitioner reports rather than empirical research. While the four models for TDR certification are all based on the concept of risk, literature in this area tends to rely on the classical definition of risk rather than to engage with the social construction of risk and the notion that risk perception can be influenced by social factors.

This literature review demonstrates that the field of digital preservation and TDR certification has thus far engaged with the concept of risk perception in only a cursory way. Rather, risk has been understood as a fixed concept that can be objectively known and understood, and scholarship in this field seems to assume that responses to risk will not vary between individuals. This dissertation applies theoretical framework based on the social construction of risk to examine the ways in which responses to risks vary depending on a variety of social factors that have been shown to influence perceptions of risk in the TRAC audit and certification process.

## **2.1 Social Construction of Risk**

In this section, I explore the social construction of risk, drawing from research in the areas of science and technology studies, organizational behavior, public policy, and risk analysis. This topic forms the basis for the theoretical framework of this study. Research in this area has found that people's actions are influenced by their perceptions of risk in addition to the objective and quantifiable value of a particular risk. These perceptions of risk are influenced by social factors, several of which are described in section 2.1.2 below. In fact, even when risks are known and understood people are poor at judging how to act in response to them.

I have chosen to focus on a subset of the risk literature – the social construction of risk and factors that influence risk perception – because this segment of the research explores risk as a socially constructed phenomenon rather than treating it only as a discrete, quantifiable figure. Section 2.1 will begin by discussing different definitions of risk, and then establishing a more nuanced definition of risk. This is followed by a critical review of the ways that risk perception is constructed by individuals and organizations (2.1.1). In section 2.1.2 I investigate several types of factors addressed in the research that have been hypothesized or shown to influence risk perception, and then apply those factors to the processes of risk analysis and risk mitigation in section 2.1.3. In sections 2.2 and 2.3 I will address digital preservation and risk as well as TDRs. These sections will show that existing research around digital preservation and TDRs has relied on a classical definition of risk, positioning a theoretical framework focused around risk perception and the social factors that influence perceptions of risk to make a unique contribution to the field.

### 2.1.1 Definitions: Risk and Risk Perception

A classical definition of risk comes from the Royal Society and includes elements that are common throughout the literature: “the probability that a particular adverse event occurs during a stated period of time, or results from a particular challenge. As a probability in the sense of statistical theory, risk obeys all the formal laws of combining probabilities” (Royal Society (Great Britain) & Study Group on Risk, 1983). This report defines risk as the probability of some adverse event. Risk has also been defined as “the potential for realization of unwanted, negative consequences of an event” (Rowe, 1977, p. 24). Other definitions of risk include the magnitude of the negative consequences of that adverse event: “[r]isk is the combination of the likelihood of an event and the consequences of that event” (Leveson, Dulac, Marais, & Carroll, 2009, p. 230). And taking this definition one step further, “[d]efinitions of particular risks include at least three conceptual elements: an *object* deemed to ‘pose’ the risk, a putative *harm*, and a *linkage* alleging some form of causation between the object and the harm” (Hilgartner, 1992, p. 40).

As these examples demonstrate, risk is generally described as some combination of probability and magnitude of consequence relating to a hazard or adverse event (Gardoni & Murphy, 2013; Hilgartner, 1992; Kaplan & Garrick, 1981; Leveson et al., 2009; Rowe, 1977; Slovic, 1987). The assumption is that both probability and magnitude can be reduced to measurable quantities and calculated as a numeric value (Fischhoff, 1983; Fischhoff, Hope, & Watson, 1990; Fox, Gardner, Lees, Green, & Andrews, 1981; Kaplan & Garrick, 1981; Starr, 1969). Each of these definitions includes the common elements of an adverse event or harm, and the probability that the event will happen. Cumulatively, these resources produce a definition of risk that assumes that both the probability that an adverse event will occur and the magnitude of the consequences of this adverse event are knowable, discoverable, and calculable, and that any

reasonable person would reach the same conclusion when calculating a given risk (Kaplan & Garrick, 1981). These definitions of risk rely upon the concept of a rational actor, and fail to take into account the idea that social factors might influence the way that individuals perceive risk (including probability, magnitude, and/or the object of risk):

*“The classical risk approach assumes that it is possible to define and assess risks. The assumption that risks can be objectified and calculated has met with a lot of criticism. Notions like complexity and uncertainty to characterize the risk situation have played a central role in clarifying the limits of the classical risk approach.” (van Est et al., 2012, p. 1075)*

Bayesian statistics allow for differences in risk perception in the context of incomplete information (Silver, 2012; Viscusi, 1985). However, this approach deals with the limits of our knowledge rather than the different ways in which people respond to the same information (Silver, 2012). This approach also assumes that as more information is known, predictions of risk will become more accurate.

In contrast, Renn argues that “risks are created and selected by human actors” (Renn, 2008, p. 11). And, given the fact that risks are fundamentally human creations, he defines risk perception as “the outcome of processing, assimilation and evaluation of personal experiences or information about risk by individuals or groups in society” (Renn, 2008, p. 64). Risk perception assumes that “the concept ‘risk’ means different things to different people” (Slovic, 1987, p. 283). Theories of risk perception hold that different people have differing understandings of the probability and adverse consequences of events, and that these differing understandings are the result of social, organizational, and/or political factors (Lachlan, Burke, Spence, & Griffin, 2009; Nelkin, 1989; Nickel & Vaesen, 2012; van Est et al., 2012; Wildavsky & Dake, 1990). It should also be noted that while there are many competing theories that seek to explain risk perception, it is still very much “a phenomenon in search of an explanation” (Sjöberg, 2000, p. 1).

The following sections will explore how eight factors influence risk perception: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. These eight factors are relevant for digital preservation because the ecosystem of TDRs is one in which complex factors combine under conditions of uncertainty. Multiple stakeholders across the different systems have varying levels of expertise, trust, and vulnerability that influence their perceptions of risk, and multiple overlapping organizations require that individual actors communicate both within and across groups.

## **2.1.2 Factors That Influence Risk Perception**

### **2.1.2.1 Communication: Amplification/Attenuation**

Risk perception can be influenced by the ways in which risks are communicated (Bostrom, 2014; Chung, 2011; Kasperson & Kasperson, 1996; Konheim, 1988; Lachlan et al., 2009; Renn, 1991; Renn, Burns, Kasperson, Kasperson, & Slovic, 1992). Theories of risk amplification and attenuation argue that risk is socially constructed and that, "the human experience of risk is simultaneously an experience of potential harm and the ways by which institutions and people process and interpret these threats" (Kasperson & Kasperson, 1996, p. 96). That is, information about risk is communicated in different ways and that, in turn, influences information processing and interpretation by people and organizations.

The amplification (or attenuation) of risk information can take place in a myriad of ways, and can involve many different types of actors and organizations, including media, government, political actors, scientists, or other experts (Arvai, 2007; Kasperson & Kasperson, 1996; Lachlan et al., 2009). Different individuals, if given the same risk information from the same source, will not necessarily perceive risks in the same way, and so it is important to consider both the audience and the mode of communication any time risk information is communicated (Arvai,

2007). For example, Lachlan et al. (2009) found that risk messages about Hurricane Katrina were effective for some groups but not others, based on race. The effectiveness of risk communication about Katrina influenced decisions about whether or not to evacuate ahead of the storm, and the result of this difference in how risk messages were received led African American respondents to evacuate at lower rates than Caucasian respondents. One of the factors influencing perception of risk in this case is the source of the risk communication and whether the messages are coming from information sources that are relevant to the recipient (Lachlan et al., 2009).

Furthermore, "[r]isk analysis, then, requires an approach that is capable of illuminating risk in its full complexity, is sensitive to the social settings in which risk occurs, and also recognizes that social interactions may either amplify or attenuate the signals to society about risk" (Kasperson & Kasperson, 1996, p. 96). Risk analysis or assessment processes can have unintended consequences on risk perception that arise not from the information itself but rather from the social factors surrounding the risk, the assessment process, and the ways in which all of those things are communicated. Risk perception can be amplified or attenuated as a result of the ways that technical information about risk is communicated.

This suggests that the discourse at several levels influences risk perception in digital repositories. Communications between repository staff regarding risk may result in either amplification or attenuation of risk depending on the mode of communication, the way in which the risk message is communicated, and the relationship between the source and recipient of the message. These factors may also influence the risk perceptions with regard to communication between repository staff and auditors in the TRAC process. Professional discourse around digital preservation and TDRs may also influence perceptions of risk for different types of stakeholders depending on their connections to the digital preservation community, and the ways in which the

communication norms of this community either match or diverge from the communication norms of their own professional communities.

#### **2.1.2.2 Complexity**

The second factor, complexity, has also been found to influence risk perception. Research shows that high levels of complexity in technical and social systems can make it difficult to identify probabilities, consequences, and hazards (Fischhoff, 1983; Perrow, 1999; Rijpma, 1997; van Est et al., 2012). Wilkinson (2001) argues that “[a]ny attempt to mask the complexity of the social experience of risk perception in rigid conceptual abstractions may lead us further away, rather than towards a more intimate understanding of the day-to-day reality in which people recognize and negotiate with ‘hazards’ as ‘risks’” (Wilkinson, 2001, p. 11).

Charles Perrow (1999) concludes that in complex and tightly-coupled systems, accidents are inevitable. He presents the example of a nuclear power plant, and demonstrates that complexity results in interactions between seemingly independent features (Perrow, 1999). Others, including Jos A. Rijpma (1997), have maintained that complexity may neutralize the benefits of redundancy, and also impair organizational learning. In other words, complexity introduces problems while also counteracting measures that are meant to offset those problems.

The research presented above indicates that perceptions of risk can be influenced by the complexity of technical and/or social systems. This suggests that the differing levels of complexity of digital repositories and in the organizations within which they are situated may lead to varying perceptions of risk for different stakeholders depending on their level of familiarity with different aspects of the repository. Additionally, digital repositories engage in complex tasks including the technical work of preserving digital information.



In the TRAC audit and certification process, repository staff face complexity in terms of the technical work of digital preservation, and also in terms of organizational and economic factors relating to the repository. Auditors may assess each repository differently depending on the complexity of the repository environment, the complexity of their own organization, and/or the complexity of the audit process itself.

### **2.1.2.3 Expertise**

The third factor that has been found to influence perceptions of risk is expertise. A great deal of literature about risk perception focuses on the differences between experts and lay people (Douglas & Wildavsky, 1982; Hilgartner, 1992; Kasperson & Kasperson, 1996; Konheim, 1988; Perrow, 1999; Slovic, 1987; Tversky & Kahneman, 1974; E. Vaughan & Seifert, 1992; Wildavsky & Dake, 1990; Wynne, 1992). It was once thought that experts had more accurate understandings of risk because they had greater levels of knowledge about the factors that contribute to risk and that their understanding of risk was more objective and/or rational than the understanding that lay people had of risk (Otway, 1992; Starr, 1969; Wynne, 1992).

However, as researchers have developed a deeper understanding of risk perception and the ways that experts and lay people differ in their perception of risks, research has come to support the idea that risk assessment and management efforts should include both perspectives:

*"[p]erhaps the most important message from this research is that there is wisdom as well as error in public attitudes and perceptions. Lay people sometimes lack certain information about hazards. However, their basic conceptualization of risk is much richer than that of the experts and reflects legitimate concerns that are typically omitted from expert risk assessments. As a result, risk communication and risk management efforts are destined to fail unless they are structured as a two-way process. Each side, expert and public, has something valid to contribute. Each side must respect the insights and intelligence of the other." (Slovic, 1987, p. 285)*

Research has found that the distinction between experts and lay people is misguided, and that individuals with different experiences and types of knowledge bring different types of expertise to bear on assessments of risk (Pidgeon, 1998). Rather than a clear division between experts and laypeople, a digital repository consists of people with varying levels of expertise in different aspects of the repository. Broadly speaking, repositories generally consist of people with administrative expertise, digital preservation expertise, and IT expertise, and each of these types of people are involved in the TRAC audit and certification process. Each of these types of people have deep, focused knowledge in some areas of repository management but may be considered laypeople with regard to others. This knowledge, paired with a lack of similar expertise in other areas of repository management, has the potential to influence perceptions of risk by opening their eyes to some types of risk and closing them to others. With regard to the TRAC process, these varying levels of expertise may influence the process differently depending on how involved each person is in the process, and on how much they rely on and trust the expertise of others.

Expertise as described in this section above and trust are closely linked with regard to risk perception. With regard to risk management for digital repositories, and TRAC audit and certification specifically, participation is required from a variety of people who have different types of expertise. Trust in the expertise and knowledge of others is necessary in order to complete the documentation required, and research has shown that trust is a factor that can influence perceptions of risk.

#### **2.1.2.4 Organizations**

Fourth, organizations are “both centres for processing and handling risks and potential producers and exporters of risk” (Hutter & Power, 2005b, p. 1). Bridget Hutter and Michael

Power (2005b) argue that risk analysis and risk management are both activities that take place within organizations, and that these activities rely on social constructions of risk knowledge that are framed within the structure of the organization.

Vaughan provides two examples of organizations shaping risk perception – the National Aeronautics and Space Administration (NASA) and the National Air Transportation System (NATS) (D. Vaughan, 1996, 2005). In these examples, Vaughan argues that decisions about the acceptability of risk are shaped by the culture of the organizations in which these decisions take place. For these two organizations, levels of uncertainty play a significant role in determining acceptable levels of risk, albeit in distinctly different ways. NATS operates in a highly standardized environment in which error and mistake are not tolerated, while NASA operates with high levels of uncertainty, normalizing deviance and creating an environment that is accepting of anomalies (D. Vaughan, 2005). Whether risk is acceptable or not, both NASA and NATS have cultures that work to ensure that individuals within the organization have a shared understanding of risk, which is promoted partly through the enforcement of formal operating procedures. In both cases, this organizational understanding of risk is critical for carrying out the everyday activities of the organization.

An alternative perspective on how organizations shape risk perception is that rather than constructing a shared perception of risk for all members, individuals within an organization will perceive risk differently depending on their roles (Hutter, 2005). This view holds that, “[w]hile some risks will be common to everyone in an organization and understood in broadly similar ways, other risks may be differentially experienced and managed” (Hutter, 2005, p. 67).

Understandings of risk “tend to be situated so that how we see, what we see, how we interpret what we see, and our ability to respond are all to some extent determined by our organizational

locus” (Hutter, 2005, p. 73). This view of risk perception corresponds to the section above, which argues that vulnerability and privilege can affect risk perception – because different roles within an organization have varying amounts of power, control, vulnerability, and exposure to risk.

For TRAC certification, each repository is a separate organization that may shape perceptions of risk for its members in different ways. Auditors also belong to organizations that may shape their perceptions of risk in ways that differ from the repositories that they are assessing. It is also possible that individuals or groups within organizations will have varying perceptions of risk based on their position in the organization, in a manner similar to the influence of expertise or trust as described in sections 2.1.2.3 and 2.1.2.6 in this chapter. Position within an organization is, in some cases, related to expertise because people are likely to be situated within groups based on their expertise (e.g., IT, digital preservation, etc.). However, lines of communication within organizations do not always follow these functional lines. Organizations with matrixed reporting structures, for example, may influence perceptions of risk in different ways than more traditional or siloed organizations.

#### **2.1.2.5 Political Culture**

Political culture, the fifth social factor that influences risk perception, argues that national context influences how risks are defined (Beck, 1992; Jasanoff, 1986; Parthasarathy, 2007). Karl Dake (1991) argues that “mental models of risk are not solely matters of individual cognition, but also correspond to *worldviews* entailing deeply held beliefs and values regarding society, its functioning, and its potential fate” (Dake, 1991, p. 62). This argument is based on the assumption that individuals exist within social, cultural, and political spheres, that they perceive risks within those contexts, and that their perceptions of risks are influenced by those contexts (Dake, 1991).

Indeed, Shobita Parthasarathy (2007) argues that political culture shapes practices and artifacts in ways that vary across political boundaries, and that the differences among political cultures can explain some of the challenges to transnational technology transfer. This may account for the slow uptake of the TRAC standard outside of the United States.

Perceptions of risk are shaped not only by the political culture within which individuals exist, but also by their place or role within that culture (Beck, 1992). For Ulrich Beck (1992, 1999), relative power within political culture influences perceptions of risk in that individuals, organizations, or groups with greater power are able to reduce their exposure to risk. Much like theories of vulnerability described in section 2.1.2.8 above, risk perception depends in part on whether an individual has control over their own level of exposure to risk.

In addition to the fact that political culture shapes perceptions of risk, Sheila Jasanoff (1998) contends that, “Theories of risk perception are inherently political because they carry within them implicit understandings about how to organize and implement policies for managing risk ... people's attitudes toward risk partly reflect their feelings of power, or lack thereof, in relation to the sources of risk” (Jasanoff, 1998, p. 93). Individuals perceive risks within their own cultural and political context, and it is within this same context that decisions about how to respond to those risks are formulated and implemented. In the case of digital preservation, repository managers may be influenced in their perceptions of risk by political events such as the de-funding of heritage organizations on a national scale (CBC News, 2012).

The view that political culture influences risk perception is relevant for the TRAC audit and certification process because of the strong ties between political culture and policymaking (Jasanoff, 1998; Parthasarathy, 2004; Wohlers, 2010). Thus far, all of the CRL TRAC certified repositories are in North America (two in Canada, four in the United States). It is possible that

perceptions of risk within the audit and certification process will vary across national boundaries. This may arise as differences in repository management between the Canadian and American repositories, and may also surface in the interaction between Canadian and American auditors.

#### **2.1.2.6 Trust**

Sixth, trust influences risk perception. The notion that risk assessment and management processes are limited by the focus of experts whose perception of risk lacks valuable information that can only be provided by non-experts is reflected in Wynne's case study of radioactivity and Cumbrian sheep farmers (Wynne, 1992). Wynne found that the judgment of officials relied on generalizable scientific principles and failed to consider the important local knowledge of the farmers (Wynne, 1992; Yearley, 2000). This case provides an example of different types of knowledge contributing meaningfully to a risk management process, including the traditionally recognized expertise of the scientists, as well as the local expertise of the farmers. It also demonstrates the importance of trust for risk perception. In the case of the farmers, the "insensitivity of official scientific agencies to the complexity of any and all local circumstances and to the (usually) corresponding richness of local knowledge" negatively affected the relationship of the farmers to the officials, resulting in a lack of trust (Yearley, 2000, p. 106). Yearly's research is also an example of actors with different types of expertise. The scientists and farmers each relied on different types of knowledge, based on their own expertise. The gap in knowledge resulted in their inability to trust one another.

Wynne argued that "public experiences of risks, risk communications, or any other scientific information is never, and can never be, a purely intellectual process, about reception of knowledge *per se*" (Wynne, 1992, p. 281). Rather, "the trustworthiness and credibility of the social institutions concerned are basic to people's definitions of risks" (Wynne, 1992, p. 300).

Information about risks cannot be separated from its context. In the case of the Cumbrian farmers, risk information was communicated by officials who demonstrated a lack of knowledge about, and respect for, particularities about the location and groups involved. The farmers disagreed about the assessment of risk and resulting policies because they did not trust the individuals and institutions involved, and the external officials had failed to establish trust with the farmers in part by failing to respect their knowledge and expertise. The two groups – scientists and farmers – had very different perceptions of risk, both of which were informed by the norms around knowledge production within their communities.

Similarly, Aaron Wildavsky and Karl Dake (1990) found that “the great struggles over the perceived dangers of technology in our time are essentially about trust and distrust of societal institutions, that is, about cultural conflict” and that “risk perceptions and preferences are predictable given individual differences in cultural biases” (Wildavsky & Dake, 1990, pp. 56, 57). Dorothy Nelkin argues that not only does trust affect perceptions of risk, but that trust (or mistrust) can be a guiding factor in how risk is defined: “[d]efining risk can become a way of explaining the failure of existing political or social relationships, of voicing mistrust, of delegating blame” (Nelkin, 1989, p. 98). Findings from Brian Wynne (1992), Wildavsky & Dake (1990), and Nelkin (1989) indicate that relationships exist between individuals and institutions, and among people with different types of expertise, and that trust is an important factor in these relationships. Lack of trust both influences and can be influenced by risk perception, and can impact efforts to assess and manage risk.

TDR certification, and TRAC in particular, is all about demonstrating that repositories can be trusted. The goal of the TRAC checklist is for repository staff to demonstrate their ability to manage risk in a number of areas. As with expertise, the management of digital repositories

depends upon different stakeholders within the organization having trust in others. The push for transparency in the TRAC audit and certification process is meant to foster trust among repository staff and with external stakeholders such as auditors. Perceptions of risk may be influenced by the levels of trust across individuals and groups within each repository. In the audit and certification process for TDRs, trust between the repository and the auditors may also influence perceptions of risk. This suggests that perceptions of risk for digital repositories may be influenced in part by relationships among people and groups both within and across organizations.

#### **2.1.2.7 Uncertainty**

Seventh, uncertainty has been identified as a factor that influences risk perception, “in many circumstances, it is not self-evident to define what the hazards, their probabilities, and the consequences precisely are” (van Est et al., 2012, p. 1076). Scholars have characterized risk calculations that take place under conditions of both speculation and ignorance as representing uncertainty about either the probability or magnitude of consequences of an event (Starr, 2003; van Est et al., 2012). More recently, scholars have argued that the dichotomy between probability and magnitude is flawed and that it is more productive to talk about risks themselves as uncertain rather than uncertainty in particular elements of risk: “current risk assessment is mostly future-oriented. The basis for risk assessment, therefore, has shifted from probability, based on experience in the past, to possibility, based on expectations about the future” (van Est et al., 2012, p. 1077). In this view, probability and magnitude cannot really be separated when considering uncertainty for risk. Rather, these elements combine to make risks themselves uncertain.



The research presented here indicates that perceptions of risk can be influenced by the existence and recognition of uncertainty. This factor may affect different types of people in different ways, as levels of expertise or knowledge about particular events, systems, or risks can influence the degree of uncertainty that a given person perceives. Digital repositories in particular face uncertainty with regard to organizational stability and funding.

Uncertainty may influence the TRAC audit and certification process because the repository staff members who take part in the process are likely to have varying levels of expertise and knowledge about the repository. Additionally, auditors are likely to have varying levels of knowledge about the repository and may experience uncertainty based on their own knowledge and expertise about activities relating to digital preservation.

#### **2.1.2.8 Vulnerability**

Eighth, risk perception may also be influenced by factors that heighten vulnerability or risk exposure, such as gender or socioeconomic status. Research in this area argues that lived experience, including exposure or vulnerability to risk, can influence risk perception (Konheim, 1988; Olofsson et al., 2014). In an article seeking to expand risk perception research in the area of gender differences, Hitchcock makes the point that people who benefit less from high-risk technology, and who lack control over their own exposure to those technologies, live in a more dangerous or risky world than people who benefit from these technologies or who are able to limit their own exposure to them (J. L. Hitchcock, 2001). Another study which focused on gender, race, and risk perception with regard to environmental risks concluded that, “perhaps women and nonwhite men see the world as more dangerous because they benefit less from many of its technologies and institutions, and because they have less power and control” (Flynn, Slovic, & Mertz, 1994, p. 1107).

These researchers argue that groups of people who lack privilege may face greater exposure to risks, and that this exposure may be thought of as independent of will or choice. When viewed alongside Chauncey Starr's (1969) findings that "the public is willing to accept 'voluntary' risks roughly 1,000 times greater than 'involuntary' risks," this suggests that risk perception will fluctuate as privilege and the ability to control one's risk exposure varies (Starr, 1969, p. 1237). Individuals who lack the ability to control their environment, and who do not benefit from the sources of risk, may perceive greater risk in any given situation than individuals who occupy positions of relatively greater privilege. In other words, choice matters and the ability to choose one's risk exposure influences how much risk a person is willing to accept in any given situation.

Michelle Murphy's "Sick Building Syndrome and the Problem of Uncertainty" (2006) provides an in-depth examination of how office workers – mainly women – were exposed to risks. This work is not framed as an examination of risk perception, but rather as an examination of the production of uncertainty. Nevertheless, Murphy argues that "[s]ociety is set up to protect the privileged from toxic events" and goes on to explain that government agencies, such as the Environmental Protection Agency, generated "suspensions of perception" with regard to risks from harmful chemical exposure (Murphy, 2006, pp. 111, 124). In this case study we see an example of a vulnerable group being exposed to harm/risk, and of a powerful group refusing to recognize that exposure. The level of vulnerability influenced the risk perception of each group, with more vulnerable individuals perceiving greater harm or risk associated with the risk object than those in positions of greater power.

Understanding how vulnerability can influence risk perception is important for TDRs because different repositories and stakeholders have varying levels of vulnerability to external

factors (based on location, financial resources, etc.), which influence the risk perception of stakeholders in the TDR system. For example, decisions about the geographic location of primary and backup storage sites are made for many different reasons, some practical and some political. Repository staff may have different perceptions of risk depending on their own involvement in the selection of sites. Similarly, repositories that lack economic security (e.g., heavy reliance on ‘soft money’) may also have different perceptions of risk than repositories with more financial stability. For the TRAC audit and certification process, perceptions of risk may vary among repository staff depending on their awareness of these vulnerabilities. Perceptions of risk may also vary between repository staff and auditors, as awareness of vulnerability (auditors) and exposure to risk (repository staff) do not influence perceptions of risk to the same degree (Starr, 1969).

The elements described above represent eight social factors that have the potential to influence perceptions of risk within the TRAC audit and certification process: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. Uncertainty and complexity may influence perceptions of risk held by repository staff and auditors because of the complex nature of digital repositories and the challenges associated with understanding the probability or magnitude of future events. Repository staff and auditors may have expertise in a variety of areas and these varying levels of expertise may heighten awareness of some risks and diminish awareness of others. The TRAC process requires input from people across many different functions and roles in a repository, and perceptions of risk may be influenced by levels of trust across those different people. Trust between repository staff and auditors may also influence perceptions of risk. Perceptions of risk may vary across repository staff and auditors depending on their awareness of, and exposure to, vulnerabilities. And finally,

communications among repository staff and auditors may amplify or attenuate perceptions of risk depending on how the message is communicated.

### **2.1.3 Risk Assessment & Management**

Risk assessment and risk management processes both influence and are influenced by risk perception (Hutter & Power, 2005b; Pearce, Russell, & Griffiths, 1981; van Est et al., 2012). The goal of understanding risk is to find ways to manage or mitigate risk for individuals and groups. In order to do that, we must first assess risk. Risk assessment “is about assessing (technical) risks” and providing “input into the risk management process, that is, the political decision-making process about how to deal with risk” (van Est et al., 2012, p. 1070). Risk assessment under the classical interpretation of risk was a (relatively) simple matter of calculating the probability and magnitude of an adverse event. As understandings of risk expanded to include greater amounts of uncertainty, social factors, and varying perceptions, the process of risk analysis became increasingly complicated.

Risk assessment and management processes may include people with different types of expertise. Research by van Est et al. (2012) regarding the unique perspective that each type of stakeholder has about risk are reflected in studies addressing the need for broad participation in risk analysis and management processes (Pearce et al., 1981; Renn, 1999). Thus, a risk assessment conducted solely by technical experts will be narrowly-focused and will possibly lack broad contextual information that is necessary for implementation of a full risk management plan (Wynne, 1992). For TDRs this means that at the repository level, the development of risk management plans should include input from individuals throughout the organization. For example a risk management plan created by the technologists might focus too specifically on the technology and not on the financial or organizational issues impacting technology.

Most formal risk assessment and management takes place within organizations; “[i]f selection of risk is a matter of social organization, the management of risk is an organizational problem” (Douglas & Wildavsky, 1982, p. 198). The ways that the organization and/or its members might share or have differing perceptions of risk can influence the outcome of the risk assessment or management activities (Hutter, 2005; Hutter & Power, 2005b; Renn, 2008; D. Vaughan, 1996, 2005). Similarly, outcomes can be affected by whether the individuals conducting the assessment are a part of the organization or if they are outsiders, as this will also influence their likelihood of sharing the risk perception of the organization being assessed (D. Vaughan, 1996).

Organizations that are highly complex, or that operate with high stakes, may have rigid risk assessment processes (Leveson et al., 2009; Perrow, 1999; D. Vaughan, 1996, 2005). One example of an organization that is highly complex and faces high stakes is NASA. Vaughan’s examination of the way that risk was managed around the Challenger disaster illustrates how risk assessment and management take place in a highly complex organization, and shows that problems can arise when stakeholder groups do not have a shared understanding of the activities that are being assessed (D. Vaughan, 1996).

Multiple organizations may be involved in a particular repository. For example, Chronopolis consists of a partnership between three organizations: UC San Diego Library (UCSDL), National Center for Atmospheric Research (NCAR), and University of Maryland Institute for Advanced Computer Studies (UMIACS) (“About Chronopolis,” 2016). In the case of the Challenger disaster, cultural differences between NASA and Morton Thiokol, Inc. on one hand, and the auditors who assessed the incident after the fact, resulted in interpretations that differed greatly (Feynman, 1988; D. Vaughan, 1996). For repositories with organizational

partnerships such as Chronopolis, cultural differences among repository staff may contribute to the complexity of the risk assessment process, both among members of the organization and for external auditors.

## **2.2 Digital Preservation and Risk**

In this section I explore the topic of risk in the digital preservation literature. I begin by establishing a definition of digital preservation in order to establish the boundaries of the scholarship that will be reviewed in section 2.2.1. Section 2.2.2 critically reviews the ways that research in digital preservation has addressed risk and risk perception. What we will see in this literature is that risk is a foundational element of digital preservation (Conway, 1996). Despite this strong relationship between risk and digital preservation, research in this area has tended to treat *risk* as synonymous with *vulnerability* or *threat*, and focuses on identifying or classifying those risks, vulnerabilities, and/or threats for the purposes of risk assessment and management. Digital preservation policies tend to focus on this approach to risk, attending to those that have the potential to cause harm.

### **2.2.1 Definition: Digital Preservation**

In order to understand TDRs, it is necessary to also examine digital preservation, as the goal of the TDR is to preserve digital information. “Digital preservation is the conservation of all digital materials, whether they were born digital . . . or whether they have been digitized from analog materials” (Routhier Perry, 2014). Digital preservation consists of those actions that ensure the viability and authenticity of digital objects over time, “[d]igital preservation can encompass a range of activities, from simple replication and storage to more complex transformation, depending on the assessed value and risk to the target content” (S. Hitchcock et al., 2007, p. 1). Francine Berman states that preservation actions are those “undertaken to ensure

the long-term viability and availability of the authoritative nature of digital material. Preservation actions should ensure the material remains authentic, reliable, and usable while its integrity is maintained; such actions include validation, assigning preservation metadata, assigning representation information, and ensuring acceptable data structures and file formats” (Berman, 2008, p. 55). These definitions all depict digital preservation as an ongoing activity, or set of activities, that ensures the viability of digital objects over time, and the definition of digital preservation that I will employ here includes such concepts as digital curation and digital stewardship (Lazorchak, 2011).

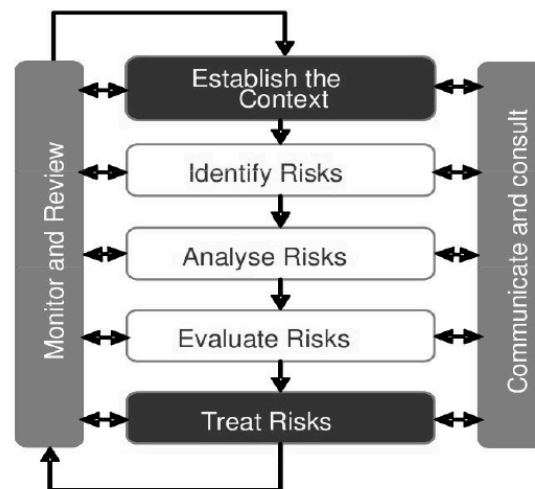
Risk is also an element of digital preservation. Conway (1996) characterized digital preservation as an ongoing process of risk management. Some definitions characterize digital preservation as a type of risk assessment (Conway, 1996; Vermaaten et al., 2012) or risk management (Barateiro et al., 2010), and others describe digital preservation as consisting of actions or practices that include risk assessment and/or risk management (Barateiro et al., 2010; Ross & McHugh, 2006a, 2006b; Strodl, Becker, Neumayer, & Rauber, 2007). Across these characterizations, there is broad acceptance of the notion that the preservation of digital objects and risk assessment or risk management practices are related.

Since digital preservation is the primary goal of TDRs (Garrett & Waters, 1996), and the concept of digital preservation must be considered within a context of risk, it follows that research about TDRs should consider the concept of risk.

### **2.2.2 Risk in the Digital Preservation Literature**

Despite the prominence of the word *risk* in literature about digital preservation and TDRs, I was able to locate only one article that gave a definition of the term. Barateiro et al. (2010), uses a classical definition of risk: “where risk is defined as the combination of the

probability of an event and its consequences” (Barateiro et al., 2010, p. 6). The authors explain that the goal of risk management is to provide a model for risk assessment and to find ways address risk for specific activities and objects (see Figure 2 below).



**Figure 2: Risk Management Process (Barateiro, et al., 2010)**

Barateiro et al.’s model (2010) includes activities shown to influence risk perception, such as communication, but makes no mention of how varying perceptions of risk might impact this process. This is all the more interesting because they propose a risk management based approach to the design and assessment of digital preservation environments. A risk management solution which assumes that everyone will perceive the same risks and will respond to them in the same way fails to take into account the fact that there are many factors which can influence perceptions of risk which, in turn, can influence the ways in which they respond to the risks that they perceive. Additionally, a risk management approach that doesn’t take into account the external factors that influence digital preservation processes presents an unrealistic view of the ways in which digital repositories carry out their work.



This is not surprising, as none of the digital preservation literature reviewed for this dissertation presented any acknowledgment that social factors might influence perceptions of risk, or any references to the large body of risk and risk perception research from other disciplines such as psychology or science and technology studies. More importantly these social factors have not been considered in the context of TDR audits.

Other discussions of risk management and digital preservation include the creation and testing of models to determine what evidence would be sufficient in order to demonstrate that a repository can effectively manage risk (Ross & McHugh, 2006a, 2006b). This course of research is directly related to TDRs as the goal is to understand what evidence repositories should present during an audit for certification as a TDR. Barateiro et al. (2010) have criticized this approach arguing that risk assessment and management should be built into the digital preservation systems themselves rather than applied after the fact as assessment criteria.

Despite this criticism, digital preservation literature continues to treat *risk* as synonymous with *vulnerability* or *threat*, and focuses on identifying or classifying those risks, vulnerabilities, and/or threats for the purposes of risk assessment and management. The terms vulnerability and/or threat are used here in a way that is more consistent with a computing/systems approach than with the risk perception research discussed above. Vermaaten et al. have identified three approaches to digital preservation and threats: typologies of threats associated with a single aspect of digital preservation (e.g., ingest), case studies of the application of threat typologies, and general digital preservation threat typologies (Vermaaten et al., 2012, p. 3). Aside from Vermaaten et al. (2012), articles about threats associated with particular aspects of digital preservation are numerous (De Vorsey & McKinney, 2010; Lawrence et al., 2000; Wright, Miller, & Addis, 2009).

Literature that addresses risk by developing typologies (Barateiro et al., 2010; Clifton, 2005; Dappert, 2009) has significant limitations for digital preservation and risk management. While the framework of treating risk as a statistical expectation is not inconsistent with managing and responding to risks, it does not necessarily address the fact that those responses will also depend on how those risks are perceived. Typologies of risk assume that different people will perceive risks in the same way and fail to take into account factors that can influence perceptions of risk. McHugh (2012) argued that, “[c]riteria lists are by their nature somewhat ineffective in illustrating the interconnectedness of infrastructural facets that can increase or limit risk exposure in various ways” (McHugh, 2012, p. 3). Digital repositories are complex systems; a listing of criteria or threats, even if it is supported by research, is not sufficient to capture the various factors that can influence risk exposure. This perspective, that complexity in systems can affect risk exposure, is supported by research in the area of risk management (Leveson et al., 2009; Perrow, 1999; Rijkma, 1997).

Given our understanding of the ways that complexity can influence risk and risk perception, and the interconnectedness of the technical and infrastructural elements of systems such as TDRs, it follows that research about risk and digital preservation that emphasizes listing threats and/or criteria does not capture the full picture. In addition to this type of approach, research that investigates risk and TDRs in the area of digital preservation while seeking to understand risk as a complex system that includes technical and social elements is needed.

### **2.3 Trusted Digital Repositories**

In this section I explore the topic of trustworthy digital repositories (TDRs). I begin with a critical examination of literature regarding digital repositories and trust (2.3.1), which is followed by a review of three different standards for TDRs in section 2.3.2. This is important

when examining risk perception because the audit and certification processes described below are carried out by groups of people whose actions are influenced by their perceptions of risk.

### **2.3.1 Understanding Digital Repositories and Trust**

In 1996 the *Report of the Task Force on Archiving of Digital Information* described digital archives “strictly in functional terms as repositories of digital information that are collectively responsible for ensuring, through the exercise of various migration strategies, the integrity and long-term accessibility of the nation’s social, economic, cultural and intellectual heritage instantiated in digital form” (Garrett & Waters, 1996, p. 8). This report found, among other things, that “long-term preservation of digital information on a scale adequate for the demands of future research and scholarship will require a deep infrastructure capable of supporting a distributed system of digital archives” and that this infrastructure will require a sufficient number of “trusted organizations” with the capabilities needed to carry out activities related to preservation and access for digital information (Garrett & Waters, 1996, p. 40). Garrett and Waters (1996) also called for a certification process in order to create a climate of trust with regard to the long-term preservation of digital information.

The Garrett and Waters report highlighted several ideas that have been expanded upon since 1996. Namely, long-term preservation of digital information is now understood to require ongoing work, preservation will rely on infrastructures that did not exist at the time, and the concept of trust will take a prominent place in questions of digital preservation and digital repositories.

Trust in organizations has been studied in many different disciplines, such as organizational studies and management (e.g., Bryce, 2007; Pirson & Malhotra, 2011; Rousseau, Sitkin, Burt, & Camerer, 1998; Tyler & Kramer, 1995), sociology (e.g., Fenton, Passey, &

Hems, 1999), psychology (e.g., Kramer, 1999), and business (e.g., Nolan, 2007). In the area of digital preservation, Day (2008) draws upon these different areas of research, arguing that “trust is at least partly about participants accepting a level of vulnerability in exchange for certain perceived benefits” and also that “trust is developmental, as it usually builds up as organizations work together over time” (Day, 2008, p. 21). In an article discussing the development of TDRs, Dale and Gore (2010) argue that, “long-term digital preservation [can] not occur in a vacuum but instead exist[s] within a larger organizational ecosystem” (Dale & Gore, 2010, p. 17). These definitions of trust for digital repositories include factors identified as relating to risk perception, such as organizations and vulnerability although the authors fail to make this connection. This suggests that trust and trustworthiness can vary depending on the risk perception of the person or group.

In the domain of digital preservation, discussions about trust have focused on online environments (e.g., Berman, Kozbial, McDonald, & Schottlaender, 2008; Colati & Colati, 2009; Corritore, Kracher, & Wiedenbeck, 2003; De Santis, Scannapieco, & Catarci, 2003; Dryden, 2011; Kelton, Fleischmann, & Wallace, 2008; MacNeil, 2000; Mutula, 2011; Yoon, 2014) , and also on the development of criteria for the evaluation of digital repositories (e.g., Becker & Rauber, 2011; Day, 2008; RLG-NARA Digital Repository Certification Task Force, 2007). This scholarship relies upon the idea that “[t]rust is instrumental for the preservation of digital media” (Hart & Liu, 2003, p. 95). Given that trust is necessary for the preservation of digital information, and that this preservation takes place in digital repositories, we must consider how trust is conceptualized for digital repositories. Ronald Jantz and Michael Giarlo (2007) note that digital repositories are unique amongst the various types of digital organizations (such as for-profit businesses conducting e-commerce), and as such “for the digital repository, trust involves

scholarship, authenticity, reliability, and persistence over time and has little relationship to immediate financial rewards” (Jantz & Giarlo, 2007, p. 197).

This line of inquiry – scholarship about digital preservation and trust – leads to the concept of the trustworthy digital repository. While there are several different standards for TDRs, common elements of these standards include a commitment to providing reliable long-term access to digital information, and the desirability that repositories demonstrate that they are able to do so (Consultative Committee for Space Data Systems, 2012; Dale & Gore, 2010; Data Seal of Approval, 2014a; Keitel, 2012; McHugh et al., 2008). Trustworthy digital repositories are “trusted, reliable, sustainable digital repositories capable of handling the range of materials held by large and small research institutions” (Dale & Gore, 2010, p. 16). The Garrett and Waters report (1996) was the first to establish a framework of attributes and responsibilities for trustworthy digital repositories, and several standards for TDRs draw upon this framework (Consultative Committee for Space Data Systems, 2012; Dale & Gore, 2010; McHugh et al., 2008; nestor Certification Working Group, 2013).

### **2.3.2 Audit Processes for Trustworthy Digital Repositories**

In response to the need for digital repositories to demonstrate trustworthiness with regard to their ability to preserve and provide access to digital information over time, major organizations in the United States and Europe have developed processes to assess and certify digital repositories as trustworthy. Four prominent certification processes are TRAC, nestor, Data Seal of Approval (DSA), and CoreTrustSeal (CTS). TRAC and nestor are both based on the ISO 16363 standard, and are administered by external auditors, while DSA and CTS audits are conducted by members of the DSA and/or CTS community.

In the context of these processes for certification, *trustworthy* indicates that the repository has demonstrated that it has a strong technological base for a robust digital curation lifecycle, a viable organizational framework including a succession plan should the organization cease to exist, and a firm economic plan that supports all curation activities (RLG-OCLC Working Group on Digital Archive Attributes, 2002). Yet, we do not know if these audit and certification processes have plausible routes to being effective in the long run, nor how successfully this normative practice has been communicated by the creators of the standards to the auditors and to repository leaders.

Each of the processes described below approaches risk probabilistically as something that can be identified and managed, and assumes that trustworthiness can be determined through a process of risk assessment. This will be discussed in greater detail below.

#### **2.3.2.1 ISO 16363: Trustworthy Repositories Audit and Certification (TRAC) & ISO 16919: Requirements For Bodies Providing Audit And Certification Of Candidate Trustworthy Digital Repositories (PTAB)**

TRAC is a certification process whose origins can be traced to a 2002 report from the Research Libraries Group (RLG) and Online Computer Library Center, Inc. (OCLC) and is partially a reaction to the Garrett and Waters report's emphasis on trust in the repository as an important element of digital preservation (RLG-OCLC Working Group on Digital Archive Attributes, 2002). This report expresses a need for a formal certification process to assess digital repositories, and in response to this call the Trustworthy Repositories Audit and Certification: Criteria Checklist was developed (RLG-NARA Digital Repository Certification Task Force, 2007). TRAC was approved as an ISO standard in 2012 (ISO 16363) (Consultative Committee for Space Data Systems, 2012). "TRAC describes approximately 90 characteristics that must be

demonstrable by repositories that aspire to a certifiable, trustworthy status” (McHugh et al., 2008, p. 132).

The TRAC standard was developed jointly by the digital preservation community and the space data research community, specifically by representatives from the Research Libraries Group (RLG), the National Archives and Records Administration (NARA), the Center for Research Libraries (CRL), and the Consultative Committee for Space Data Standards (CCSDS) (Yakel, 2007). The groups that currently maintain the standard and administer the audits to determine certification are distinct and separate.

The TRAC standard explains that a TDR “will understand threats to and risks within its systems” (Consultative Committee for Space Data Systems, 2012, p. 19). TRAC also discusses financial risks, infrastructure risks, and security risks, and calls for repositories to identify preservation risks and provide strategies for dealing with them (Consultative Committee for Space Data Systems, 2012). The standard treats risks as concrete and identifiable and addressable, suggesting an alignment with the classical definition of risk as a quantifiable value based on the probability and consequences of a negative event, as described in section 2.1 above. The standard provides lists of suggested forms of evidence to document different items listed in the checklist, but speaks primarily to specific types of policy documents rather than defining high-level concepts such as risk and does not provide any indication that perceptions of risk may vary between and/or among the standards developers, auditors, and repository leaders.

In addition to the standard and the organizations that contributed to its development, the Primary Trustworthy Digital Repository Authorisation Body (PTAB) is another part of the sociotechnical system that comprises TRAC. The organization that aims to play “a major role in training auditors and repository managers” in the implementation of ISO 14721 (the OAIS

model), ISO 16363 (TRAC), and ISO 16919 (a standard that specifies the competencies and requirements on auditing bodies) (ISO-PTAB, 2011). PTAB appears to have arisen in response to the need for auditors who administer the TRAC standard to be trained in a consistent way. It also represents a concerted effort to broaden the base of auditors for the TRAC standard beyond the United States and CRL, as the training sessions for PTAB have been focused on international audiences, with previous training sessions having taken place in Greenwich, UK, and Pasadena, CA, and future courses scheduled for Den Haag, CERN, Washington D.C., Italy, India, and China (“PTAB Courses,” 2015).

Currently (as of July 2018) four comprehensive repositories have been certified as trustworthy digital repositories by CRL: Portico, HathiTrust, Chronopolis, and Canadiana.org (Center for Research Libraries, 2010, 2011, 2012; Free, 2011; Kirchhoff et al., 2010). Two other repositories, Scholars Portal and CLOCKSS, have been certified as trustworthy repositories by CRL for solely their e-journal content, meaning that only a portion of each repository is certified (Center for Research Libraries, 2013, 2014). One repository has been certified as trustworthy by PTAB: The National Cultural AudioVisual Archives (NCAA), hosted by the Indira Gandhi National Centre for the Arts Audio/Visual Repository (Giaretta, 2018; Primary Trustworthy Digital Repository Authorisation Body Ltd., 2018). Additionally, the U.S. Government Printing Office (GPO) has declared an intent to become TRAC certified by CRL, and has announced that they have begun work to prepare for the audit and certification process (Federal Depository Library Program, 2014).



### **2.3.2.2 DIN 31644: nestor Seal for Trustworthy Digital Archives**

The nestor Seal for Trustworthy Digital Archives (nestor) is an assessment for digital repositories that is based on the DIN 31644 standard<sup>1</sup> (Keitel, 2012; nestor Certification Working Group, 2013). The nestor process can be used by repositories for self-assessment and can also be used to obtain certification (Keitel, 2012). Like TRAC, the nestor assessment process is based on the OAIS model (Consultative Committee for Space Data Systems, 2012a, 2012; Keitel, 2012) and can be applied to a variety of different types of repositories. The nestor working group was founded in 2002, and extended certification began in 2013 after a pilot certification of the German National Library (Keitel, 2014). As of July 2018, four repositories have received nestor certification: TIB Technische Informationsbibliothek Hanover, ZBW Leibniz-Informationszentrum Wirtschaft, Data Archiving and Networked Services (DANS), and Deutsche Nationalbibliothek (nestor-Siegel, 2018).

The nestor checklist refers to risk only in the criteria relating to security and integrity, with explanatory notes instructing repository leaders to provide risk management plans as evidence during an audit (nestor Working Group Trusted Repositories - Certification, 2009). As with TRAC, nestor appears to treat the concept of risk as identifiable. nestor is described as falling in the middle of a spectrum of certification options, as it “is more elaborate and its results offer greater accuracy than that of a simple self-assessment, yet it is less elaborate and is less accurate than an intensive audit conducted by external experts as part of a formal certification procedure” (nestor Certification Working Group, 2013, p. 3).

---

<sup>1</sup> The full text of DIN 31644 is available in German only at this time.

### **2.3.2.3 Data Seal of Approval**

The Data Seal of Approval (DSA) assessment consists of sixteen guidelines, which “recognize that responsibility for archival quality data is shared amongst three groups: producers for the quality of the research data themselves, the repository for the quality of data storage and availability, and consumers for the quality of data use” (Ball, 2010, p. 31). The DSA documentation describes itself as a “trust-based accreditation” but does not position itself as a risk assessment or management process (Data Seal of Approval Board, 2013b, p. 4).

DSA is administered by Data Archiving and Networked Services (DANS), and was established by the Royal Netherlands Academy of Arts and Sciences and the Netherlands Organization for Scientific Research in 2005. DSA has evolved to become an international certification process that is administered and managed by an international board. The first edition of the DSA guidelines was presented and officially handed over to an international board in 2009 (Data Seal of Approval, 2014a). The DSA guidelines are a basic set of criteria that were designed to facilitate awareness and serve as a first step toward a “heavier” assessment and certification (Harmsen, 2008, p. 2).

There are currently (as of July 2018) 86 repositories that are DSA certified. The Data Seal of Approval is a seal which can be displayed on a repository’s website indefinitely once acquired (Data Seal of Approval, 2016). DANS describes the DSA certification as a step on the way to other, more onerous certifications; the audit process involves the repository conducting a self-audit which is then reviewed by a Data Seal auditor, a process that is relatively lightweight and low-cost in comparison to the TRAC certification process (Data Seal of Approval, 2014b).

The DSA certification system is quite large, with several stakeholder groups in addition to the standards developers and repositories described above. The DSA is governed by the DSA

General Assembly (Data Seal of Approval, 2015), which is the group that elects the DSA Board. The Board is drawn from the General Assembly, and is responsible for overseeing DSA business and communicating with the General Assembly and DSA Community (i.e., organizations with DSA-certified repositories). The DSA system also contains a group of peer reviewers who are responsible for assessing DSA applications. For DSA, peer reviewers are required to have gone through at least one successful DSA self audit. In this case, the stakeholder groups of auditors and repository managers have substantial overlap, as being a repository manager is a prerequisite to qualifying as an auditor.

The concept of risk is foundational to DSA certification. While the DSA documentation focuses on the concepts of quality and sustainability throughout, implicit in this discussion is the notion that in order to support the sustainability of digital information and to ensure the quality of that digital information, a repository must be able to identify and mitigate risk (Data Seal of Approval Board, 2013a). The audit process in which members of the DSA community serve as auditors once their own repository has achieved certification assumes that all members of the community will both understand and respond to risk in the same way.

In July 2017 DSA suspended new applications and shifted applicants to the newly formed CoreTrustSeal certification process. As DSA-certified repositories come up for certification renewal, they will be directed to CoreTrustSeal, which is described further in section 2.3.2.4 immediately below.

#### **2.3.2.4 CoreTrustSeal**

A new repository certification organization launched in 2017: CoreTrustSeal (CoreTrustSeal, 2017). This new organization has replaced both DSA and the ICSU World Data System (WDS) organizations by merging the requirements of each into one new certification, the

CoreTrustSeal Data Repository certification (Dillo & De Leeuw, 2018). Focused on data repositories, this certification is managed under the umbrella of the Research Data Alliance (RDA), and CTS certification will supersede both DSA and WDS certifications from 2018 onward (CoreTrustSeal, 2017). As of July 2018 there are 28 CTS certified repositories, and 109 repositories with legacy DSA, WDS, or DSA and WDS certifications that are expected to renew certification with CTS within the next few years (“Core Certified Repositories,” 2017; Dillo & De Leeuw, 2018).

Certification criteria for CTS fall into three categories: (1) organizational infrastructure, (2) digital object management, and (3) technology (CoreTrustSeal, 2016). These criteria are in accordance with other repository certifications such as nestor, DRAMBORA, and TRAC (Dillo & De Leeuw, 2018, p. 165). As with DSA, CTS audits are conducted by a review board consisting of individuals from CTS certified repositories as well as CTS board members (CoreTrustSeal, 2018).

#### **2.3.2.5 The European Framework**

TRAC, DSA (now CoreTrustSeal), and nestor represent a spectrum of certification options for digital repositories in which the scale of certification options is intentional. Together, these three standards now form the European Framework for Audit and Certification of Digital Repositories (Consultative Committee for Space Data Systems, Data Seal of Approval Board, & DIN Working Group “Trusted Archives - Certification,” n.d.; Giaretta et al., 2011). The framework came together after each of the certification processes had already been established, and was formed through a memorandum of understanding signed by representatives from each organization (Giaretta, Harmsen, & Keitel, 2010).

This framework places repository assessments on a scale from least rigorous (DSA/CTS) to most (TRAC), with level of rigor being determined by such factors as amount of evidence required to support claims, and the intensity of the audit conducted (Giaretta et al., 2011; nestor Working Group Trusted Repositories - Certification, 2009). This framework allows the three certifications, which are largely based on the same foundational concepts and documents (e.g., OAIS), to exist alongside one another without having to compete directly. Each certification has a niche, and meets a slightly different need. These different levels of rigor also provide repositories both a lower cost path to certification for repositories, as well as a stepwise process to move from DSA/CTS certification to full TRAC certification.

Literature addressing TDR audit and certification processes is lacking in empirical research. Most of the scholarship in this area consists of case studies by practitioners and descriptive overviews of the standards. The research methods I describe in Chapter Chapter 3: are intended to address the need for empirical research in this area.

## **2.4 Conclusion**

Risk has been understood by digital preservationists primarily through a classic lens, focusing on the concept as a calculable, discoverable value, in the area of digital preservation. Research regarding risk perception has shown that risk is a socially constructed concept that can be influenced by social and organizational factors. Since standards for TDR audit and certification processes are based on the concept of risk, it follows that the outcomes of these audits may vary depending on the perceptions of risk held by different stakeholders.

Research examined in each of the areas described in this chapter is mainly qualitative, with the exception of risk perception, which consists of both qualitative and quantitative research, including laboratory experiments. Case studies and model building are the two most

commonly represented research methods, appearing across the digital preservation and risk literature. Other research methods used include historical studies, ethnographies, surveys, interview-based qualitative research, and laboratory experiments. Research about risk perception in a laboratory setting has produced consistent, replicable results that can make strong causal claims about factors affecting risk perception, but these findings are not necessarily applicable outside of a laboratory setting because much of the real-world context is stripped away in order to conduct this type of research (Bernard, 2012; Creswell, 2009). On the other hand, ethnographic research and in-depth case studies may produce results that examine risk perception in a real-world setting, allowing the researcher to understand how social context can affect risk perception, but the findings from these types of studies are not necessarily transferrable to other cases (Denzin, 1997; Yin, 2003).

Research methods across all areas tend toward case studies and model building. In each instance, the researchers are focusing narrowly on a particular system, organization, or phenomenon and either studying it in-depth, or using it to build a model that can theoretically be applied to other situations. While many of the models developed were carefully planned, and some were presented with evidence of pilot tests, I have found very little evidence of researchers testing models that were developed by others. The literature indicates that researchers are aware of the models being developed by others, as many models are presented with extensive reviews of the state of the field (Vermaaten et al., 2012), but that the development of models happens within silos of collaborators (Barateiro et al., 2011, 2012, 2010; Barateiro & Borbinha, 2012; McHugh et al., 2008; Ross & McHugh, 2006b, 2006a). In contrast to the literature reviewed regarding risk perception, the models developed in digital preservation have been tested and interrogated through other methods of research, including case studies and surveys. However, the

models are often based on technical reports or practitioner experience rather than empirical research (Dobratz & Schoger, 2007). The models in this discipline (such as OAIS, TRAC, nestor, DSA, and CTS) are well developed and have typically been accepted by the community.

While all of the research areas include scholarship that presents literature reviews or overviews of a field, the TDR and digital preservation areas in particular lack empirical research. Literature that is also represented prominently, but that does not rise to the level of research, includes policy documents, standards documentation, technical reports, literature reviews, practitioner reports, and position papers. The case studies in this space are often conducted by practitioners who are describing their own experiences (e.g., Kirchhoff, Fenton, Orphan, & Morrissey, 2010; Minor et al., 2010; The Data Seal of Approval Board, 2011). While case study research from the perspective of a member of the organization being studied is not inherently problematic, there are issues that can arise as a result of the researcher being a stakeholder in the case. Robert Yin (2003) has noted that researchers conducting case studies must understand the issues that they are studying beforehand, and as such are prone to developing preconceived positions. As someone who is not only familiar with the issues, but who has a stake in the organization, a case study researcher studying her own organization is likely to be susceptible to this type of bias. None of the case study research examined above acknowledged or addressed possible researcher bias.

The need for empirical, rigorous research about digital preservation and TDRs demonstrates a gap in the research. The research methods described in Chapter Chapter 3: below address this deficit by presenting a qualitative study that relies on a theoretical framework based on the social construction of risk and theories of risk perception in order to understand how

differences in risk perception may influence the outcomes of TDR audit and certification processes.

This study addresses the gaps identified above by examining the TRAC audit and certification process in order to understand how stakeholders in the process (i.e., standard developers, auditors, and repository staff) understand risk and what factors influence their perceptions of risk. This addresses both the need for empirical research about TDR certification as well as the need to broaden understandings of risk and risk perception with regard to digital preservation activities.



## **Chapter 3: Research Methods**

In this qualitative study, I critically examined how individuals in three target participant groups, standard developers, auditors, and repository staff, understand the concept of risk for digital repositories in the context of a TRAC audit.

This study is motivated by the following research questions:

1. How do standard developers, auditors, and repository managers conceptualize risk in the context of a TRAC audit?
2. What are the differences and similarities by which standard developers, auditors, and repository managers understand risk as it has been communicated by the TRAC standard?
  - a) In what ways do these differences and similarities become manifest in the TRAC audit process?
3. To what degree do the following eight factors which influence risk perception come into play in the audit process: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability?
  - a) In what ways and why do they emerge when staff and auditors consider risk factors articulated in the TRAC standard?
  - b) What additional factors, if any, emerge which also influence perceptions of risk in relation to the TRAC standard?

As described in section 2.3.2.1 above, ISO 16363 is based on the idea that a trustworthy repository must demonstrate that it is able to anticipate/identify, assess, manage, and mitigate risk. If the developers of the standard, the auditors who enforce the standard, and repository staff

members do not share the same perceptions of the risks that they are attempting to identify, assess, manage, and mitigate, then the outcome of the certification process lacks meaning and weight. Certification assumes that standard developers, auditors, and repository staff members share common ground with one another; the outcome of the process depends on them all sharing conceptual clarity. If they are not conceptualizing core concepts in the same way, then the outcomes of the assessments are unreliable.

In this dissertation, I examine whether standard developers, auditors, and repository staff share the same perceptions of risk in the TRAC audit process. Through interviews with standard developers, auditors, and repository staff, and document analysis of the ISO 16363 standard, repository prepared responses to the TRAC checklist, CRL certification reports, and publications written by repository staff, I examine perceptions of risk in order to gauge whether those groups share the same understanding of risk and identify factors that may influence their perceptions of risk with regard to TRAC audit and certification.

This focus on understanding how standard developers, auditors, and repository staff involved in the TRAC audit and certification process define and understand risk, and what factors influence their perceptions of risk, called for a qualitative approach (Newman & Benz, 1998). As noted in the previous chapter, research about risk perception is primarily qualitative because of the nature of the inquiry, which focuses on understanding phenomena within their cultural context and building theories based on data rather than testing hypotheses (Newman & Benz, 1998). A qualitative approach is most appropriate in this case because of the small population size (there were, at the time of data collection, six TRAC certified repositories), and because the research questions that I asked focused broadly on the perceptions of risk within the TRAC audit and certification process (Creswell, 2013). Qualitative research emphasizes “the

collection of data in a natural setting sensitive to the people and places under study, and data analysis that is inductive and establishes patterns or themes” (Creswell, 2013, p. 37).

Additionally, this is an exploratory study, which calls for qualitative methods to elicit all of the factors that may influence perceptions of risk (Creswell, 2013), rather than a study seeking to test hypotheses about risk perception or quantify the TRAC audit process (Newman & Benz, 1998).

John Creswell (2013) has identified several characteristics of qualitative research that apply to this study. First, qualitative research design embraces the idea of multiple realities, exploring the subjective experiences of all participants in the study in order to provide evidence from different perspectives. This is key for studying the multidimensional concept of risk perception. Second, qualitative research also encourages the researcher to conduct studies in the field, getting to know the participants on their own terms and in the contexts within which they usually operate. This enables a deeper understanding of the TRAC audit process. Third, qualitative research acknowledges the biases and values of the researcher as well as the participants, and allows for examination of how those actors may influence or shape the results of the research. Fourth, qualitative research allows the researcher to define important concepts as they arise through the course of research, rather than relying on rigid predetermined definitions. This fourth characteristic was particularly important for this study. I began with eight factors that have been shown to influence how individuals construct their understanding of risk, but allowed for the possibility that additional factors would emerge during data collection and analysis. Fifth and finally, qualitative research emphasizes studying phenomena within their own contexts, beginning with particular details (such as specific questions about the TRAC audit process) and then gradually moving out to generalizations.

### **3.1 Pilot Study**

In December 2015 and January 2016, I conducted a pilot study with an organization that closely resembles the target population for this dissertation. This pilot study consisted of interviews, a survey, and document analysis of an organization that participated in a TRAC test audit, but was not TRAC certified. This particular organization was selected as it was the best fit for my target population without having to use one of the six TRAC certified repositories and thus limit the options for the full dissertation.

I interviewed four individuals from this repository, including one administrator, a digital preservation manager, a digital preservation intern, and an IT manager. Each interview lasted approximately one hour, and questions focused on the timeline and logistics of the audit process, interaction with auditors, and risk perception with regard to specific areas of the TRAC audit checklist.

Each interview was followed by a survey questionnaire that asked participants to rate their perception of the levels of risk associated with specific parts of the audit checklist. The survey was introduced at the end of the interview, and the survey questionnaire was sent via email to each participant immediately following their interview. Three of the four participants completed the survey, and one failed to complete the survey after multiple reminders, with a partial response recorded.

The interviews were audio recorded and transcribed for analysis. The transcripts were analyzed using the qualitative data analysis software package NVivo. The code set used to analyze this data was developed based on concepts from the literature, themes that emerged during the interviews, and themes that arose during analysis.

The research methods described below were developed and refined based on the results of this pilot study.

### **3.2 Analytical Focus**

The analytical focus for this study was guided by the social construction of risk, and focused on eight factors that have been shown to influence perceptions of risk: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. Although these factors formed the basis for this analysis, I was mindful that additional factors could emerge through the research process.

CRL announced it would begin certifying repositories in 2006<sup>2</sup>, and the first repository to receive TRAC certification from CRL was Portico in 2010 (Center for Research Libraries, 2010). I focused on audits that occurred from 2010-2015, (Canadiana.org, Chronopolis, CLOCKSS, HathiTrust, Portico, Scholars Portal), and also on individuals who were affiliated with PTAB.

### **3.3 Data Collection & Analysis**

I employed a qualitative research design, incorporating semi-structured interviews and document analysis. Interviews were carried out with standard developers, auditors, and repository staff. These individuals were situated within organizations, but variations in perceptions of risk have been shown to exist at the individual level, even within the same organization (D. Vaughan, 1996). While this research design was influenced by case study methods (i.e., multiple methods of data collection, looking at the TRAC standard as one case, or

---

<sup>2</sup> <https://www.crl.edu/archiving-preservation/digital-archives/certification-assessment>

at the six TRAC certified repositories as comparative case studies), the analytical focus is at the individual rather than the organizational level (Yin, 2003).

Semi-structured interviewing is considered best in situations where the researcher will only be able to conduct one interview with each participant, which was the case for this study (Bernard, 2012). This type of interview allows the researcher discretion to follow leads as they arise in the course of the interview, while providing a general guide that will ensure that the data generated are reliable and comparable (Bernard, 2012, p. 182). While interviews have the advantages of allowing the researcher control over the line of questioning, the quality of data collected may vary due to the presence of the researcher, the fact that not all respondents will be equally articulate and perceptive, and the information gathered will be filtered through the views of the interviewees (Creswell, 2009). A semi-structured interview addressed some of these limitations by allowing me to probe for additional information when appropriate during the interviews (Babbie, 2010, p. 277).

I complemented the semi-structured interviews with document analysis. Documents contain text but should also be considered as situated products created in social settings (Prior, 2003). Document analysis emphasizes not only the content or text of the documents, but also the context in which a document was produced, as well as the ways in which a document functions in specific circumstances (Prior, 2003). While many of these documents were created with the intention of being made available to others (e.g., audit findings reports prepared for the general public), none were prepared for the purposes of this study. As such, some amount of social desirability bias may be present, but this bias may communicate something interesting about the dynamics between stakeholder groups.

### 3.3.1 Population and Sample

At the start of data collection for this study (2016) four comprehensive repositories had been certified by CRL as trustworthy digital repositories: Portico, HathiTrust, Chronopolis, and Canadiana.org (Center for Research Libraries, 2010, 2011, 2012, 2015). Two other repositories, Scholars Portal and CLOCKSS, had been certified as trustworthy repositories for solely their e-journal content, meaning that only a portion of each repository is certified (Center for Research Libraries, 2013, 2014). Table 1 below shows the certification date for each repository, as well as a breakdown of the audit report scores. CLOCKSS and Scholars Portal, the two repositories that have been certified as trustworthy for only part of their content have received the highest scores.

**Table 1: Overview of CRL Audit Report Scores**

	Canadiana.org	Chronopolis	CLOCKSS	HathiTrust	Portico	Scholars Portal
CRL Certification Date <sup>3</sup>	7/1/2015	3/1/2012	7/1/2014	3/1/2011	1/1/2010	2/1/2013
Organizational Infrastructure	4	3	4	2	3	3
Digital Object Management	3	4	4	3	4	4
Technologies, Technical Infrastructure, Security	4	4	5	4	4	4
Total (max possible = 15)	11	11	13	9	11	13

The population for this study consisted of: (1) standard developers (i.e., PTAB board members), (2) auditors from CRL, and (3) staff members from the six TRAC certified repositories. Interviewees were recruited through a combination of convenience and snowball sampling. I was able to identify all of the PTAB board members, and CRL auditors and

---

<sup>3</sup> Determined by publication date of Certification Report: <https://www.crl.edu/archiving-preservation/digital-archives/certification-assessment>

certification advisory panel members, as well as at least one staff member from each repository who was involved in the TRAC audit process. I also asked interviewees to recommend others for interviews, and through this process was able to recruit three to five participants from each TRAC certified repository who participated in the audit in some manner.

**Table 2: Overview of Interviewees**

	Roles (n)			Total
	Administration	Digital Preservation	IT	
Standard Developers	0	8	3	11
Auditors	4	6	0	10
Repository Staff	9	6	6	21
Total	13	20	9	42

### **3.3.1.1 Standard Developers**

I conducted 11 interviews with PTAB Board Members and participants from the training workshops that they conducted. The PTAB Board consists of individuals who participated in the development of the ISO 16363 standard. In addition to developing the standard, they also conducted six text audits with the ISO 16363 checklist across Europe and North America. Some members of the PTAB Board have conducted training workshops around the ISO 16363 standard. Included in this study are both instructors and participants from those training workshops. In June 2017 the PTAB organization became accredited to conduct audits for the ISO 16363 standard.

Of the 11 PTAB members and trainees interviewed, eight described their current or most recent role as being digital preservation-focused, and three as being IT-focused. Two were retired at the time of their interviews, two were professors, six described their current positions as very senior, although they mostly described their roles as technical rather than managerial or administrative, and one was in a mid-level digital preservation role. Ten of the PTAB



interviewees described their involvement in repository certification as an opportunity to leverage their experience and expertise to give back to the digital preservation community.

Education information was available for eight of the 11 interviewees in this group. Four had doctoral degrees (in information science, information management, history, and theoretical physics), three had master's degrees (in information science, computer science, and history), and one had a bachelor's degree (humanities). Eight were located in the United States, and three in Europe.

### **3.3.1.2 Auditors**

I conducted interviews with 10 auditors and audit advisory board members from CRL who participated in repository audits. Auditors were those individuals employed by CRL, and advisory board members were individuals from CRL member organizations who were invited to participate in the audit process by reviewing documentation submitted by repositories and making recommendations to the auditors. The term auditors will refer to both auditors and audit advisory board members throughout this study. Three of the auditors were CRL staff members or interns, and two were consultants, and five held management positions in academic libraries. Interviewees in this group tended to frame their participation in the TRAC audit process in terms of their leadership experience in academic libraries. Of those interviewed, six described their current roles as being digital preservation-focused, and four described their current roles as administratively-focused. Education information was available for eight of the 10 interviewees in this group. All eight had a master's degree in library and/or information science. Six were located in the United States, and four in Canada.

### **3.3.1.3 Repository Staff Members from TRAC Certified Repositories**

My data collection strategy focused on recruiting three to five participants from each of the six repositories completing the audit and certification process. I also analyzed publicly available documentation including certification reports created by CRL and made available via their website.

From the pilot study as well as previous research and a review of the literature I was able to determine that three primary types of repository staff are involved in the TRAC audit process: repository administrators, digital preservation staff, and IT staff (Frank & Yakel, 2013). These three distinct groups have different types of expertise within the repositories, and these functions correspond to different sections of the TRAC checklist, which includes items relating to organizational infrastructure, digital object management, and infrastructure and security risk management (Consultative Committee for Space Data Systems, 2012).

In the pilot study, I was able to interview four repository staff members: a high-level administrator, two digital preservation professionals, and an IT manager. These four people represented the functional areas that were most heavily involved in the TRAC audit, as reported by the administrator, who was the head of the repository and lead manager of the audit preparation process. As the person who oversaw the entire process, including coordinating all of the work within the repository and also managing all communication with the auditors, the administrator was uniquely situated to understand the involvement of repository staff in the process. She was also able to provide documentation and correspondence from the time period of the audit to support this distribution of labor. These individuals/roles were each able to discuss different parts of the process, and each functional area provided a unique perspective about the repository's risk exposure. In total, they covered all parts of the TRAC self-study and actual

audit process. Interviews with the two digital preservation professionals addressed overlapping content, but since they were at the repository at different times they each had unique experiences to discuss.

Education information was available for all of the interviewees in this group. One had a doctorate degree (in mechanical engineering), 16 had a master's degree (in library and/or information science, and atmospheric and/or oceanic science), three had a bachelor's degree (in computer science, mathematics, and English), and one completed some coursework but did not have a college degree.

#### *3.3.1.3.1 Canadiana.org*

Founded in 1978, Canadiana.org is a nonprofit coalition of 40 Canadian memory institutions, and since 2005 Canadiana.org has taken a leadership role in digital preservation across Canada. In addition to preserving and providing access to digital resources, Canadiana.org is an aggregator of metadata from partner organizations and enables search across their collections. Available collections include 65 million pages in total as of March 2016 (Canadiana.org, 2015). Canadiana.org's mission statement states that through the many partnerships with Canadian memory institutions, the organization is able to "spearhead digital preservation in Canada" (Canadiana.org, 2015). The organization has a volunteer Board of Directors with 10 members who serve two-year terms and is an independent nonprofit organization that is able to engage in fundraising activities independent of its partner organizations. Canadiana.org is a membership-financed organization with a tiered member dues

system. Financial information, including tax returns, for Canadiana.org is available via the Canada Revenue Agency for 2011-2015.<sup>4</sup> Canadiana.org became TRAC certified in 2015.

As an organization that focuses on both preservation and access, Canadiana.org manages relationships with people and organizations that contribute data to the repository as well as users who wish to access and use the data that the repository provides. The web of relationships that the repository manages with the memory institutions that compose the coalition, as well as partnerships with other Canadian memory institutions, and users introduces many possible sources of complexity and uncertainty. Complexity and uncertainty are factors that can influence risk perception, and repository staff, members of the Board of Directors, and people from other organizations that are affiliated with Canadiana.org may have different types of expertise, which is a factor that can influence perceptions of risk, and may have had some bearing on the TRAC audit and certification process. Similarly, trust among all of these different people and groups may be complicated in ways that could influence their perceptions of risk. Canadiana.org represents the interests of a wide array of memory institutions, and the people who manage both Canadiana.org as well as its partner organizations are likely to experience vulnerability in different ways, which may influence the ways that they perceive and manage risk for the repository. Awareness of, or exposure to, the financial management of the repository is one area where people are likely to experience differing perceptions of risk. Communication through the organization and with TRAC auditors may influence perceptions of risk, either amplifying or attenuating those perceptions, depending on a number of factors including information source, mode of communication, and the relationships between repository staff and auditors.

---

<sup>4</sup> <http://www.cra-arc.gc.ca/ebci/haip/srch/t3010form23sched6-eng.action?b=118833425RR0001&fpe=2015-03-31&n=Canadiana.org&r=http%3A%2F%2Fwww.cra-arc%E2%80%A6>

### 3.3.1.3.2 *Chronopolis*

The Chronopolis digital preservation network was originally funded by the Library of Congress (“About Chronopolis,” 2016). Chronopolis is now managed by three organizations: University of California, San Diego Library (UCSDL), National Center for Atmospheric Research (NCAR), and University of Maryland Institute for Advanced Computer Studies (UMIACS). Chronopolis is also a node in the Digital Preservation Network (DPN), managing ingest and replication for the entire network. Chronopolis focuses on preservation rather than access, and “provides services for the long-term preservation and curation of America's digital holdings” (“Chronopolis Homepage,” 2016). Financial information for Chronopolis is not publicly available, but the organization does provide pricing information for organizations wishing to sign up for the service, which indicates that Chronopolis is trying to recoup costs through a fee for service model (“Chronopolis Pricing,” 2016). Chronopolis became TRAC certified in 2012.

Chronopolis is a repository that focuses on preservation rather than access. Data depositors can request their own data, but the repository does not provide public access to its collections. While this simplifies the network of stakeholders with regard to repository users, the repository is managed by a network of three institutions, which introduces elements of complexity and uncertainty. NCAR in particular introduces both complexity and uncertainty to Chronopolis because the center manages classified government data and as such access to some of the documentation that a typical TRAC audit would require has some access restrictions. Chronopolis is managed by a number of people across several institutions, and it is likely that across the different institutions, functions, and departments, repository staff have different types of expertise in different areas. Experts in different areas are likely to have different perceptions

of risk, and it is likely that people with a variety of different types of expertise participated in the TRAC audit and certification process. TRAC places a high value of transparency as a way to demonstrate trustworthiness, but at least one Chronopolis institution has government-enforced information restrictions. It is possible that perceptions of risk are influenced by these limitations with regard to TRAC audit and certification. Chronopolis is managed across three geographically dispersed locations: San Diego, CA, Boulder, CO, and College Park, MD. Each of these areas faces different environmental and location-based vulnerabilities, which may influence the perceptions of risk for members of repository staff at those locations. Repository staff and auditors may also have differing perceptions of risk depending on their knowledge of the repository's finances. Reporting structures across Chronopolis as well as within each of the institutions that comprise the repository may influence risk perception for people who are at different places within those structures. And finally, communication throughout Chronopolis, as well as between repository staff and auditors may influence perceptions of risk.

#### *3.3.1.3.3 CLOCKSS*

CLOCKSS became a nonprofit organization in 2009, and the repository consists of a partnership with Stanford University and member organizations that pay a fee to participate. As with Chronopolis, CLOCKSS maintains a geographically distributed repository focused on preservation rather than access (Center for Research Libraries, 2014). CLOCKSS preserved e-journal content from publishers, and provides access to the data only if the content becomes unavailable from the publishers themselves. CLOCKSS is governed by a Board of Directors consisting of 22 members as well as an Executive Director (CLOCKSS, 2015). Financial information for CLOCKSS is not publicly available, but the repository does provide pricing information for libraries and publishers who may wish to join the network ("Contribute to

CLOCKSS,” 2015). CLOCKSS became TRAC certified in 2014 (Center for Research Libraries, 2014; CLOCKSS, 2014).

Although CLOCKSS is a distributed digital preservation network with numerous locations and participating organizations, the LOCKSS model for preservation is one in which most activity takes place at a central hub (i.e., Stanford), with minimal participation from other members (Reich & Rosenthal, 2001). While member organizations may experience uncertainty about the technical aspects of repository management, the structure of the network reduces their exposure to the complexity behind the scenes. Repository staff members are likely to be centralized at Stanford, which may reduce uncertainty and complexity with regard to repository management – in contrast to a repository such as Chronopolis, which is distributed across three locations. This model requires a high degree of trust among member organizations, given their low levels of involvement in the repository’s day-to-day operations. Alternately, auditors examining the repository for TRAC certification may have a limited view of the repository’s overall structure based on their experience communicating with staff at Stanford. Repository staff members are likely to have different types of expertise based on their roles and backgrounds, which may influence their perceptions of risk for CLOCKSS. Perceptions of risk are also likely to vary depending on how repository staff and auditors experience vulnerabilities or risk exposure with regard to factors such as the financial sustainability of CLOCKSS, and the technical security of the repository. Communication among repository staff, between repository staff and member institutions, and between repository staff and auditors may also influence perceptions of risk.

#### 3.3.1.3.4 *HathiTrust*

HathiTrust is a partnership of research institutions and libraries, including more than 100 partners worldwide (“About HathiTrust,” 2016). Founded in 2008 with 13 institutions, the repository contains digitized content from partner institutions, including content from the Google Books project (“Beyond Google Books: Getting Locally-Digitized Material into HathiTrust,” 2015). “HathiTrust is governed by a Board of Governors, composed of six members appointed by HathiTrust founding institutions, six elected from the membership at large, and the Chief Executive Officer, who serves as an ex officio, non-voting member” (“Governance,” 2016). The University of Michigan is the current host institution for the infrastructure of the repository. Funding for hardware, software, and services is provided through the University of Michigan Library (“Governance,” 2016). HathiTrust has an active mirror backup site in Indiana, which is maintained by staff at that location (“Getting Content Into HathiTrust,” 2016). Current collections as of April 2016 include 627 terabytes of data, consisting of nearly 14 million total volumes (“Statistics Information,” 2016). Like Canadiana.org, the repository focuses on both preservation and access to content. Financial information is not publicly available for HathiTrust, and the University of Michigan annual budget does not directly mention HathiTrust (U-M Office of Budget and Planning, 2016). HathiTrust became TRAC certified in 2011.

There are several possible sources of uncertainty and complexity for repository staff and TRAC auditors with regard to HathiTrust that could influence risk perception. Numerous partner organizations and data contributors may create difficulties for repository staff managing relationships as well as data for the repository. The fact that HathiTrust staff members are embedded within the University of Michigan Library may also create uncertainty about their roles or about differentiating HathiTrust from the library itself. People with different types of



expertise are likely to perceive risk differently, among repository staff as well as TRAC auditors. Trust among repository staff across primary and backup sites may influence perceptions of risk as well. Repository staff and auditors may have different degrees of exposure to risk, including awareness levels of the financial stability of the repository, and the stability of backup sites. And finally communication among repository staff, between repository staff and partner institutions, and between repository staff and auditors may also influence perceptions of risk in the TRAC audit process.

#### *3.3.1.3.5 Portico*

Founded in 2002, Portico is a not-for-profit organization that preserves electronic scholarly content including e-journals and e-books (ITHAKA, 2015a). Portico focuses primarily on preservation rather than access, although access to content that becomes unavailable via the publisher can be triggered upon request (ITHAKA, 2015d). This model is similar to that of CLOCKSS. Portico is a service of ITHAKA (other ITHAKA services include JSTOR), and is overseen by its Board of Trustees, in addition to an advisory committee of librarians and publishers (ITHAKA, 2015c). As of March 2016, the repository contained approximately 62.5 million preserved archival units (ITHAKA, 2015b). Portico is centrally managed, with participating organizations acting as customers or clients rather than partners. Information about the location of backup sites is not publicly available. Financial information is publicly available for ITHAKA, but this may or may not be useful in understanding the economic sustainability of Portico. Portico was the first repository to become TRAC certified in 2010.

Portico appears to have a more centralized, consolidated governance structure than the other repositories described here. For this reason, there may be less uncertainty and/or complexity for repository staff to contend with. As with the other repositories described here,

staff members are likely to have varying types of expertise which can influence perceptions of risk, and the same is true of auditors. Perceptions of risk among repository staff and auditors are likely to be influenced by their exposure to risk, or their experiences of vulnerability. For example, people with greater knowledge of the organization's financial stability may perceive risk in this area to be greater or smaller than those with less knowledge about this aspect of the repository. Communication among repository staff as well as between repository staff and auditors may amplify or attenuate perceptions of risk depending on factors such as mode of communication.

#### *3.3.1.3.6 Scholars Portal*

Founded in 2002, the “Scholars Portal technological infrastructure preserves and provides access to information resources collected and shared by Ontario’s 21 university libraries” (Ontario Council of University Libraries, 2014c). The work of the Ontario Council of University Libraries (OCUL) is governed by the University Librarians of the 21 member libraries (Ontario Council of University Libraries, 2014a). The Scholars Portal team appears to operate primarily from the University of Toronto (Ontario Council of University Libraries, 2016). The repository includes more than 40 million e-journal articles, 6 million e-books, as well as datasets and geospatial data (Ontario Council of University Libraries, 2014b). As with CLOCKSS, Scholars Portal is TRAC certified for only e-journal content. Financial information for Scholars Portal is not publicly available. Scholars Portal became TRAC certified in 2013.

Focusing on both preservation and access for digital information, Scholars Portal is a repository with many distinct stakeholder groups, including data contributors, users, repository staff, other staff from OCUL, and staff from the 21 member libraries. This suggests a high level of complexity as well as many possible sources of uncertainty for both repository staff and

auditors, which may influence perceptions of risk. Repository staff members are likely to have different types of expertise, which can also influence perceptions of risk. Scholars Portal has made all of their TRAC documentation available via the website<sup>5</sup> with some information available only upon request. This degree of transparency is very much in the spirit of TRAC, which emphasizes transparency as a measure of trustworthiness. In a model similar to HathiTrust, Scholars Portal coordinates across many institutions but is located primarily within one (i.e., University of Toronto). It is likely that information about the economic sustainability of the repository is tied to the University of Toronto in a similar manner, which may influence perceptions of risk for repository staff and auditors depending on their understanding of the organization's finances.

The six repositories that have been TRAC certified represent a range of governance models, with varying degrees of centralization. Funding models across the repositories vary with regard to the amount of information available. In each case, it will be important to ask for additional information about repository funding and financial sustainability in order to understand both repository staff and auditor perceptions of risk with regard to this important aspect of TRAC. Certified repositories also vary in their focus on preservation, access, or both. Repositories that seek to both preserve and provide access to information manage a more complex range of stakeholders, and also a more complex technological environment. The different types of expertise required for repository management varies depending on many of the factors described above, and this will influence the number and type of staff that each repository requires. These factors, along with many others, are likely to influence perceptions of risk for repository staff at the repositories as well as auditors assessing each for TRAC certification.

---

<sup>5</sup> <https://spotdocs.scholarsportal.info/display/OAIS/Document+Checklist>

### 3.3.2 Interviews

Semi-structured interviewing is considered best in situations where the researcher will only be able to conduct one interview with each participant. This type of interview allows the researcher discretion to follow leads as they arise in the course of the interview, while providing a general guide that will ensure that the data generated are reliable and comparable (Bernard, 2012, p. 182). While interviews have the advantages of allowing the researcher control over the line of questioning, the quality of data collected may vary due to the presence of the researcher, the fact that not all respondents will be equally articulate and perceptive, and the information gathered will be filtered through the views of the interviewees (Creswell, 2009). A semi-structured interview addressed some of these limitations by allowing me to probe for additional information when appropriate during the interviews (Babbie, 2010, p. 277). This was particularly important as I experienced some difficulty during the pilot study in eliciting responses that uncovered factors influencing risk perception. As a result, the interviews for this dissertation were more structured in order to keep participants on topic. Each interview was recorded and transcribed for analysis.

The first half of the interviews focused on a vignette, which was sent to all participants ahead of the interview (see Appendix 5). This vignette consisted of a repository description that I generated based on the repository site profiles in section **Error! Reference source not found.** above and the requirements described in the TRAC standard. This was not a description of an existing repository, but rather a fictional repository that shared some characteristics with the existing TRAC certified repositories. Participants were asked to discuss the vignette, identify possible sources of risk for the repository described therein, and suggest ways to address or mitigate those sources of risk. The vignette provided common ground for comparing risk

perception across participants, as the simplicity of a scenario can help to “identify, clarify, and disentangle the complexities of real-world processes” (Hughes, 2004). Notably, they can be helpful as an interview strategy when participants are highly visible and/or identifiable within their community, as standards developers, auditors, and staff members from TRAC certified repositories were likely to be (Gubrium & Holstein, 2001). Because vignettes “contribute toward understanding people's perceptions, beliefs, attitudes, and behavior” but do not necessarily allow generalization to understanding real life (Hughes, 2004), participants were also asked to discuss their own experiences.

The second half of the interviews included questions about each participant’s individual experiences either as a standard developer, auditor, or repository staff member in order to understand perceptions of risk in relation to particular repositories. Participants were asked questions that focused on each of the eight factors from my model for the social construction of risk, as well as questions about their role in the audit process. Interviews lasted approximately one hour.

In addition to this two-part interview the repository staff member who was responsible for coordinating the audit process and acted as the main point of contact for the auditors was asked additional questions about the motivation for, and logistics of, the audit and certification process. In my pilot study this was the repository manager, and this generally held true for the TRAC certified repositories where the audit process was managed by either the repository manager, or a digital preservation specialist. Only one person at each repository was asked to go through this longer interview, as I found through my pilot study that the other stakeholders did not have the same high-level view of the process and asking those questions of each interviewee did not yield additional information beyond what the administrator was able to provide.

Additionally, I wanted to avoid focusing too much on the logistics of the process and instead spend the interview asking questions that focused on how interviewees constructed their understanding of risk.

The pilot study that I conducted included a survey, which was administered to participants immediately following their interview. I decided not to include a survey as a data collection method here for several reasons. First, I found that it was difficult to obtain complete responses from all of the interviewees. In fact, only three of the four participants completed the survey, even with multiple reminders. Second, it was impractical to administer the survey as a part of the interview because I spoke with some participants in person and others remotely, making it difficult to ensure that all would have taken the survey under the same conditions. Third, it would have added a substantial amount of time to each interview, and I was already asking for a significant time commitment from each participant. Fourth, the sample for this study is too small to have any statistical power, and as such a survey with Likert scale questions would not yield particularly useful information. Asking these questions during a qualitative interview allowed me to discuss these questions in greater depth with participants, and to probe further in order to understand the reasoning behind their responses to the questions. And finally, having realized that I would not need to ask each interviewee about the logistics of the audit and certification process, I chose instead to focus the interviews more narrowly on the social construction of risk. Using the survey questions as a guide, I covered this subject matter during the course of the interview, which allowed me to ask for clarification about responses, and to help ground the questions themselves during the interview.

### **3.3.2.1 Testing Interview Protocol**

I conducted a pilot test of my interview protocol with six subject matter experts as recommended by Anderson and Gerbing (1991). These subject matter experts were representative of the main population of interest for this study (i.e., repository staff), allowing this pilot to serve as a test of the content validity. Straub et al. define content validity as “the degree to which items in an instrument reflect the content universe to which the instrument will be generalized” (Straub, Boudreau, & Gefen, 2004, p. 424). This pilot consisted of cognitive walkthroughs with the interviews in which participants were asked to use the concurrent think-aloud technique to verbalize their thoughts while answering the interview questions (Groves et al., 2009). They were also asked to reflect on both the interview questions and the vignette upon completion of the interview. Pilot test interviews were audio recorded.

### **3.3.2.2 Analysis of Interview Data**

Interviews were audio recorded and transcribed. They were then coded and analyzed using NVivo,<sup>6</sup> a qualitative data analysis software package. I employed an open coding approach, beginning with an initial set of codes based on my review of concepts from the literature, themes that emerged from the pilot study, and themes that arose during the interviews (Charmaz, 2006; Saldaña, 2015a). These codes were descriptive, analytic, and thematic. Descriptive codes identified points where participants discussed concepts such as interaction between repository staff and auditors, specific types of evidence prepared for a TRAC audit, or challenges encountered during the audit process, for example. Analytic codes included the eight factors that I identified as influencing risk perception: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. Analytic coding moves

---

<sup>6</sup> <http://www.qsrinternational.com/product>

beyond descriptive coding and allows the researcher to “transcend the local and particular of the study to more abstract or generalizable concepts” (Saldaña, 2015a, p. 120). During analysis I revisited the interview data to develop additional thematic codes as well (Saldaña, 2015b). (See Appendix 6).

This approach to qualitative data analysis resembles the technique described by Matthew B. Miles and A. Michael Huberman (1994). It is not grounded theory; rather, I began with a code set developed through my own work as well as a critical review of research literature in relevant areas (e.g., risk perception, digital preservation), but also kept an open mind in order to modify the code set as needed in response to themes that emerged during data collection and analysis.

Using the code set that I developed, I coded the interview transcripts in two groups: (1) auditors and standard developers, and (2) repository staff members. For each group of interviews, I enlisted the help of a second coder. We worked independently coding the same transcript, in order to ensure that my application of the code set was consistent and reliable. Using Scott’s Pi, a statistic measuring interrater reliability for coding textual data (Holsti, 1969; Scott, 1955), we achieved a score of 0.711 for the repository staff, and 0.719 for the auditors and standard developers. This step demonstrated that the data I collected, the code set that I created, and the data analysis that I conducted were both comprehensible to, and usable by, other researchers. This exercise also provided assurance of the reliability of subsequent data analysis.

### **3.3.3 Document Analysis**

I conducted document analysis comparing the text of the ISO16363 standard against the certification reports provided by CRL, the evidence provided by each repository in support of their audit, and publicly available documentation from each repository, where available (see 3.3.3.1 below for more information about available documentation). This analysis focused



primarily on those checklist items that have been identified as affecting risk, but considered the checklist and certification reports in their entirety as risk and risk management are foundational concepts that underlie the entire checklist.

Documents are products that are produced by people within social contexts, using technologies. Analyzing documents can provide information about the processes and circumstances through which they were created, and also about their intended audience (Prior, 2003). The comparative analysis of these documents looked within each audit to compare the evidence provided by repository staff to the response from auditors, and also across repositories to compare both the type and amount of evidence provided in response to each checklist item, and to compare auditor responses to this evidence and finally the certification scores assigned by CRL.

Drawing from Diane Vaughan's (1996) research methods, I conducted a preliminary round of document analysis with all publicly available documentation ahead of the interviews. This helped to provide guidance for the interviews, allowing me to focus the discussion around those parts of the audit and certification process that emphasize risk.

### **3.3.3.1 Available Documentation**

Document analysis included audit reports, repository prepared responses to the TRAC checklist, and publications written by repository staff. CRL certification reports detailing the results of each audit were publicly available via CRL. Four repositories have made the documentation that they presented to auditors available via their websites (CLOCKSS, HathiTrust, Portico, Scholars Portal). Canadiana.org has not made their entire checklist available, but provided a selection of preservation policies. Chronopolis has not made any TRAC

documentation publicly available. A list of publications written by repository staff was available via each repository website.

**Table 3: Summary of Available Documentation, By Repository**

Repository	CRL Audit Report	Repository Prepared Responses to TRAC Checklist	Staff Publications
Canadiana.org	✓		✓
Chronopolis	✓		✓
CLOCKSS	✓	✓	✓
HathiTrust	✓	✓	✓
Portico	✓	✓	✓
Scholars Portal	✓	✓	✓

Analysis of these documents, as well as publications by repository staff, provided additional perspectives about the TRAC certification process, as well as information about the repositories more broadly. Document analysis helped to mitigate problems of memory and recall arose during the interviews (Sudman, Bradburn, & Schwarz, 1996), as many of the documents were been produced at the time of the audit and provided information about the social and organizational context in which the audit was conducted (Prior, 2003). In several instances, interviewees had difficulty recalling dates and/or timeframes for particular events relating to specific audits. I was able to provide information about the dates of the audit from the CRL certification reports and repository publications, which they then used to find additional information about their own experiences (e.g., looking up old emails, finding events in their calendars, etc.).

### **3.4 Limitations**

Through the pilot study I ran into difficulty when participants were unable to recall events surrounding the audit, which had taken place nearly ten years prior to the interviews. The more time that has passed since an event, the greater the likelihood that a person will have

difficulty remembering it (Sudman et al., 1996; Tourangeau, Rips, & Rasinski, 2000). Including documentation and publications written by repository staff helped to mitigate this limitation. While all of the repositories in my study population have been through their audit more recently than the site in the pilot, I found that many had similar difficulties with memory and recall, most notably those from the earliest audits. Also, rationalization and sensemaking happen over time. I addressed this challenge by including links to each repository's TRAC certification report in the interview request emails, and by suggesting that participants may want to refer back to their own notes, documents, emails, or calendars either before and/or during the interview.

Audit and certification processes for trustworthy digital repositories are a relatively new phenomenon and the population that I examined in this study is small. Social desirability effects likely arose during interviews both within and across repositories, as well as among auditors and standards developers, due to the small size of the community (Bernard, 2012, p. 205). Other response effects included the expectancy effect, inaccuracy of self-reporting, and the deference effect (Bernard, 2012). In addition to being small, the population for this study consisted of individuals who were primarily based in the United States. Maintaining the anonymity of the small number of participants who were located outside of the United States. limited the analysis that I was able to conduct. For example, potential sources of risk and/or factors from the model for the social construction of risk that emphasized political and/or legal issues were likely to reveal the nationality and/or location of participants, thereby making them identifiable within their professional community.

In this chapter I have described the qualitative research methods that I employed for this examination of the social construction of risk in the TRAC audit and certification process. In the next chapter, I present my findings. Chapter Chapter 4: begins with an examination of how

individuals understand risk in the TRAC audit process. This is followed by an analysis of potential sources of risk that interviewees identified, and concludes with an examination of two specific aspects of TRAC certification in order to understand which factors from the theoretical model for the social construction of risk in digital preservation are present in the TRAC audit process.

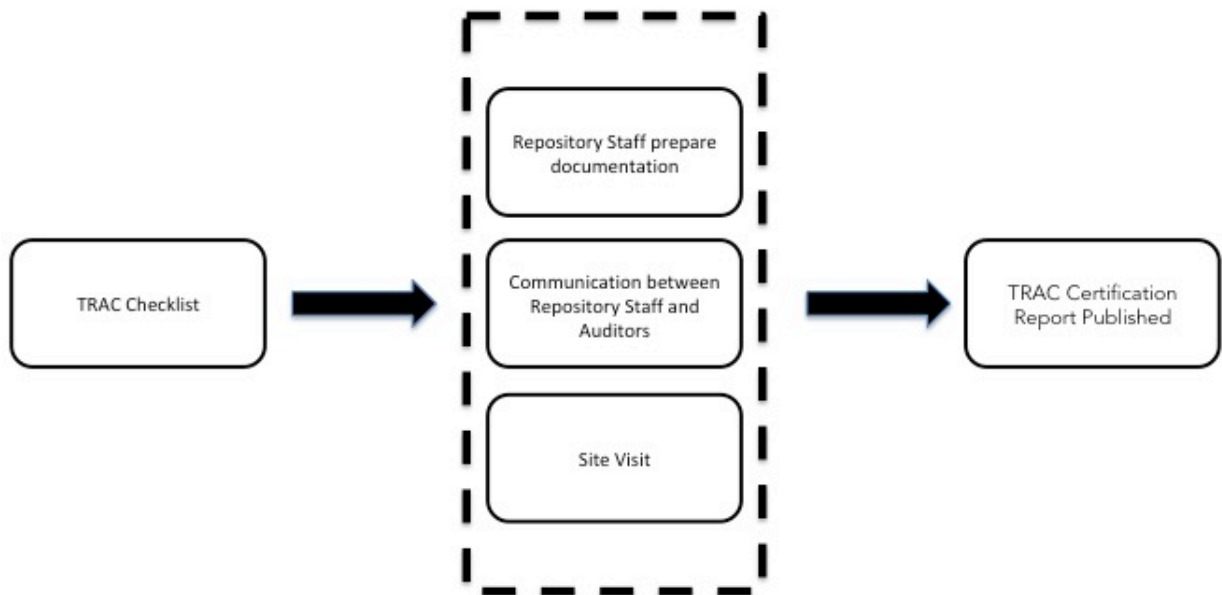
## **Chapter 4: Findings**

### **4.1 Introduction**

In this chapter I examine how standard developers, auditors, and repository staff involved in the TRAC audit process understand the concept of risk for digital preservation, including similarities and differences in their understandings of the concept of risk. I argue that although the digital preservation community has regarded the concept of risk as a calculable value, it is in fact socially constructed. Reliance on a classical definition of risk in the audit process includes an assumption that people will behave in rational and predictable ways in response to the same information. In terms of TRAC, we see this classical definition in the way stakeholders identify risks and mitigation strategies and assume that identification and planning is sufficient to demonstrate the ability to preserve content for the long-term. I found that while risk identification is the first step toward long-term digital preservation, it is when and how the mitigation strategies described are enacted that determines long-term preservation outcomes. Therefore, risk perception in the context of TRAC includes what is perceived as being the best or most effective risk response.

This chapter begins with an examination of risk that traces the TRAC audit process, including the developers of the standard that governs the process, the auditors who conduct the audits, and the repository staff whose repositories undergo audits. My analysis shows that these

three groups understand risk primarily in relation to specific threats or potential risks, which I have organized into five main categories: finance, organizational governance, legal, repository processes, and technical infrastructure. This is followed by an examination of two aspects of the audit process: the auditor site visit and process by which repositories maintain certification after the initial audit. My analysis focuses on the social factors that influence how standard developers, auditors, and repository staff members understand risk in these two areas, and emphasizes the importance of communication, expertise, uncertainty, and vulnerability in shaping perceptions of risk in the context of TRAC.



**Figure 3: TRAC Certification Process**

Overall the results present a nuanced picture of the TRAC audit process as one in which the actors involved agree on a classical definition of risk, but differ about whether an audit process based on this definition can determine trustworthiness with regard to long-term digital preservation. My findings demonstrate that while standard developers, auditors, and repository

staff generally share an understanding of the major sources of potential risk that face digital repositories, they disagree about whether and how these risks can be mitigated. Individuals who are removed from the day-to-day work of the repositories undergoing an audit are more likely to accept risk identification and mitigation strategies as sufficient evidence of trustworthiness while individuals at the repositories are skeptical that their performance of trustworthiness as proscribed in the audit process will translate to actual trustworthiness with regard to long-term digital preservation.

## **4.2 Risk and TRAC**

ISO 16363 defines a trustworthy digital repository as one that understands threats and risks, and that can communicate this understanding to the public in order to engender trust:

*“A trustworthy digital repository will understand threats to and risks within its systems. Constant monitoring, planning, and maintenance, as well as conscious actions and strategy implementation will be required of repositories to carry out their mission of digital preservation.” (ISO 16363, 2012)*

The standard presents a checklist of criteria against which to measure repositories. Many of these criteria are accompanied by directions to identify and describe risks and the types of documentary evidence a repository might provide to demonstrate risk mitigation. This approach treats risk as something that can be identified, and asserts that risk assessment is a necessary part of digital preservation. For example, criteria 4.3.1 states: “The repository shall have documented preservation strategies relevant to its holdings” (ISO 16363, 2012, p 52). And this item is accompanied by an example of ways that a repository can demonstrate it is meeting this requirement, “Documentation identifying each preservation risk identified and the strategy for dealing with that risk” (ISO 16363, 2012, p 52). If a repository is able to do both of those things across all areas of the checklist, they are awarded certification as a Trustworthy Digital Repository (TDR).

Interviewees discussed risk in ways that demonstrated an understanding of risk that was in line with the classical definition. For example, Repository Staff 18 explained that he did not think that his colleagues understood what risk means for digital repositories, and that while it is easy to find information about risk mitigation strategies it is more difficult to understand the probability and magnitude of consequences of a potential risk. This explanation highlighted his view of risk as calculable, but consisting of uncertain elements:

*“Do I think that large amounts of people really understand how risk is constructed and what it means? No. ... I think it’s relatively easy to get information about solutions and how things are implemented, and it’s harder to put that in a framework where you’re measuring the likelihood of it happening against the potential of it happening, and what the downsides are there, and how you tie specific numbers to that.” (Repository Staff 18)*

Similarly, Auditor 10 focused on a definition of risk that relied heavily on numbers and probability, and also noted that the TRAC standard does not provide a clear definition of risk: “You know, I think that’s the big challenge with this. It’s like, okay, so what if I don’t do this? Well, what’s the probability that you won’t? And what’s the probability that this could happen? But that isn’t really discussed a whole lot in the standard” (Auditor 10).

This view demonstrates an understanding of risk for digital preservation that assumes that it is important to understand risk as a calculable figure, despite the uncertainty of really being able to calculate the risk. Like the classical model of risk, this view is based on an underlying assumption that people are rational actors who will understand risk information in similar ways, and behave predictably in response to that risk information.

In this study I found that interviewees identified risk in terms of specific potential threats or sources of risk, which I have organized into five main categories: finance, legal, organizational governance, repository processes, and technical infrastructure.

- **Finance** risk refers to potential threats to the financial sustainability of a repository.



- **Legal** risk includes potential threats relating to rights management (e.g., copyright), and liability issues that could arise if a repository fails to sufficiently protect sensitive materials.
- **Organizational governance** risk refers to potential threats to the stability of a repository's organizational infrastructure, including internal governance structures within a repository and how the repository is positioned within larger organizations (e.g., universities, consortia, partnerships, etc.).
- **Repository processes** can be a source of potential risk to a repository, including metadata capture and/or creation, file format management, and the ingest or migration of digital content.
- **Technical infrastructure** related risk refers to potential threats to the hardware and/or software (e.g., servers, databases, etc.) that a repository relies on to manage digital content.

In each of these cases, interviewees identified potential threats or sources of risk to digital content as well as repositories as a whole. This understanding of risk for digital preservation is in line with existing studies of risk in digital preservation, which have focused on identifying specific threats and reporting about how individual repositories or technologies mitigate those risks (e.g., Barateiro, Antunes, & Borbinha, 2011; Vermaaten, Lavoie, & Caplan, 2012). In the following section, I will examine each category in greater detail.

### 4.3 Potential Sources of Risk

#### 4.3.1 Finance

Interviewees described financial uncertainty as a potential source of risk to digital repositories. They discussed the ways that loss of funding threatens both repositories and digital information, and framed their understanding of this threat in terms of long-term business planning and risk identification. While auditors and repository staff agreed with the conceptualization of risk presented by standard developers, they found that in practice repositories fell short of meeting the requirements of addressing risk as articulated the TRAC standard, but were still certified as trustworthy.

Although all perceived financial risk as significant for repositories, each group understood the risk differently. For example, Standard Developer 03 said that financial viability was a potential source of risk and vulnerability for digital repositories, because so few have managed to secure long-term funding and remain operational, “Well other than repositories that are institutionally mandated, a long-term business plan is very difficult to come by. You know, there are a few long-lived digital repositories that aren’t institutional repositories, but there aren’t many that have lasted very long. So just how do you ensure that you’ve got adequate funding over the long-term when people’s interests change so rapidly?” This explanation highlights both the importance of long-term funding for digital repositories as well as the difficulty in securing that funding without an institutional mandate.

Uncertainty about funding sources and the lack of stable long-term funding was also described as a significant source of potential risk for digital repositories by Standard Developers 01, 02, 03, 06, 07, 08, 09, and 10, who emphasized that digital preservation requires continuous funding, “Money, funding. Long-term availability of the data, depending upon – since long-term availability is dependent on containing funds for the data being preserved, for keeping the data preserved or at least continuous monitoring of the data” (StandardDeveloper\_01).

Digital repositories require ongoing maintenance and funding in a way that paper-based repositories do not, making financial sustainability an especially significant source of risk for digital repositories. One standard developer explained that the expertise required to maintain digital information over time depends on having financial resources, “Ongoing finances, because in a paper based repository you could put stuff in a box and leave it. You can’t do that in a digital repository. Depending on your network, how you have arranged either succession or partnerships that can help with ameliorate that, I mean, everything really does come down to money ...

Money buys you expertise, along with servers and stuff. If you don't have expertise, you're not going to be very good at this. You have to have money to do that” (StandardDeveloper\_10).

Among standard developers, relying on grant funding or other forms of soft money was not viewed as a sustainable long-term plan for repositories. Standard developers were highly critical of the funding model for the repository described in the vignette, which they thought left the repository vulnerable because it relied on grants for a substantial portion of its funding, “When you're running a repository, the implication is that the repository will be storing and providing access to resources over time. How that's going to be accomplished over time when there's grant funding for, it's really over a third of their support? That is a concern” (StandardDeveloper\_02).

The perspectives presented by standard developers about financial sustainability as a potential source of risk for digital repositories is reflected in the text of the standard itself, which governs the audit process. It is through this document that the standard developers developed and shaped an understanding of risk that includes threats to financial sustainability, and to set expectations about how repository staff could prove to auditors that they sufficiently identified and addressed those threats.

Indeed, the TRAC standard contains several references to financial sustainability, including sections 3.1 and 3.4 of the Organizational Infrastructure section, titled “Governance and Organizational Viability” and “Financial Sustainability” respectively. Checklist item 3.4.1 states that, “3.4.1 The repository shall have short- and long-term business planning processes in place to sustain the repository over time” (Consultative Committee for Space Data Systems, 2012, p. 30). The supporting text further explains that a repository should be able to “ensure the

viability of the repository over the period of time it has promised to provide access to its contents for its Designated Community” (Consultative Committee for Space Data Systems, 2012, p. 30).

Similarly, item 3.4.3 states that, “The repository shall have an ongoing commitment to analyze and report on financial risk, benefit, investment, and expenditure (including assets, licenses, and liabilities)” (Consultative Committee for Space Data Systems, 2012, p. 31).

Supporting text for this checklist item states that a repository should be able to provide a documented risk registry detailing potential threats and their corresponding mitigation techniques.

Although the TRAC standard does not specify what length of time would be considered long-term, Standard Developer 06 explained that repositories should be thinking about how to preserve digital content in terms of centuries rather than years or decades, “Well, your Designated Community has to include users who haven't even been born yet. You have to have a way in which you can ensure, guarantee, that you can preserve that information a half century, a full century, four centuries into the future. There's no way around that” (StandardDeveloper\_06).

Both of these criteria present financial sustainability as a matter of identifying specific threats and documenting strategies to respond to those threats, and argue that a repository must be able to demonstrate an ability to provide access to a specific community for the amount of time that they have promised. While this may seem straightforward, it presents a view of financial sustainability that rests upon several assumptions, including 1) an understanding of what constitutes short- and long-term timeframes, 2) that each repository has promised access for a specific set period of time, and 3) that identifying threats and initiating mitigation techniques is sufficient evidence that those risks will be recognized if or when they arise and that the mitigation techniques described will be successfully implemented. The view of risk reflects the

perspective of the standard developers, a group consisting of very experienced individuals with advanced technical degrees in fields such as physics and computer science, with a shared epistemic culture that emphasizes discoverable, calculable phenomena.

Following the lead of the standard developers, auditors also framed financial vulnerability as a source of potential risk for digital repositories. Auditors 01, 03, 04, 07, and 08 argued that sufficient funding was a basic requirement for repositories engaged in long-term preservation of digital content and that without adequate funding repositories would fail, “Basically, money is life in this case. If they don't have adequate funding, they're going to fail” (Auditor\_03). Auditor 02 expressed a similar position, saying that no policy could compensate for a lack of funding, “if money dries up it doesn't matter how many policies they have in place. They don't have the funding to do it” (Auditor\_02).

For Auditor 01, a digital preservation professional, the biggest potential risk facing digital repositories was financial. This interviewee explained that in his experience obtaining sufficient funding for long-term preservation has been difficult, even when a parent organization was committed to funding a repository, because of the difficulty of communicating about costs, “Money. Finances. Being able to get a firm commitment from an organization to pay the cost of storage over time can be very significant. Because well, it's something I've seen in my work a couple of times of people not quite realizing how much storage can cost, and then trying to find ways to lower that cost without your knowledge or maybe with your knowledge but just refusing to buy the things that are necessary in a lot of circumstances. Yeah, biggest risk is money and funding” (Auditor\_01).

As with standard developers, auditors tended to view financial sustainability as a problem of institutional support for the mission of long-term digital preservation. Auditor 10, for

example, said that, “the inability for institutions to commit to the costs of digital preservation” was a systemic problem in the field of digital preservation (Auditor\_10). This digital preservation professional went on to explain that the issue of funding and support for digital repositories is a potential threat to the entire field, “I think that in the history of academia there has been support for brick and mortar, paper, intellectual outputs. So libraries and research centers. But there has been a resistance, at the institution level, to support digital collections. And there has been some support from repository staff, and services, but I think we have seen a really big problem, internationally, in supporting the sole responsibility that academic institutions need to take, in terms of preserving content. And I think it's been a problem for decades, and I don't think it's been resolved” (Auditor\_10).

The auditors drew upon their experiences conducting audits as well as their own professional experiences in library administration and digital preservation, and their shared educational backgrounds in Library and Information Science, to characterize financial sustainability as an ideal that was difficult to achieve because so many repositories have relied on short term grant funding, “It's becoming less so, but a lot of the early efforts around this were grant funded, and that's great. It allows you some seed money to develop things, but you need long ... This is for long-term, and so you need organizational and financial sustainability for that mission to really be solid and strong” (Auditor\_08). Similarly, Auditor 04 explained that repositories start out with good intentions but the lack of sufficient resources over time prevents them from being able to put good policies and processes in place in order to carry out their mission of long-term digital preservation,

*“Bottom line is, it's a tremendous amount of resources required to do long-term preservation. Organizational commitment to those types of resources often waxes and wanes depending on where the organization is situated from a financial perspective. A lot of repositories start off with really good intentions and really*

*well-defined ... Well, that's actually not true, but they start off with the goal of having defined processes, workflows, and all that sort of stuff, and over time a lot of that stuff gets either dropped or the period between things like migration activities or even just repository auditing activities expands as the organizations are pressed for resources and staff. I think that's sort of the really the biggest thing. Everything else from a ... comes down to the challenge of technological change, but a lot of the technological change can be mitigated with sufficient resources. It's having the resources available to address those things.”*  
(Auditor\_04)

As with the standard developers, auditors recognized financial sustainability as a potential source of risk for digital repositories. While the standard developers focused on the importance of long-term financial sustainability, auditors focused on how difficult it has been in practice for repositories to secure long-term funding and the challenges that arise as a result. Auditors, who were tasked with both interpreting and enforcing the TRAC standard, shared the view that financial sustainability was crucial for digital repositories but emphasized how few repositories have been able to achieve the level of support that the standard sets as its ideal.

Repository staff framed financial sustainability as a potential source of risk for digital repositories generally, and for their own repositories more specifically. They shared an understanding of the importance of funding for their repositories, and the threat that loss of funding posed for both their repositories as well as their digital content, but described an environment in which resources were scarce, and shifting organizational priorities meant that they felt that their budgets were vulnerable and under constant threat, “Funding . . . people don't want to pay for preservation at all. The ones that, there's A and B, there's a perception that preservation should be cheaper than access storage, and the opposite is actually the case. It's always been the money is the hardest, most riskiest challenge” (RepositoryStaff\_04).

They recognized the need for sustainable long-term funding and the importance of being able to demonstrate and communicate their financial viability through documentation, such as long-term business plans and risk analyses in order to ensure that they were meeting the

promises made to their stakeholders. However, they described a wide gap between the ideal set forth by the standard developers and what they were able to achieve in their own institutions. Repository Staff 09 described the lack of control over her budget as a potential source of risk because it left them vulnerable to fluctuations and changes in the complex funding landscape of higher education, “the higher education field is complicated and budgets are subject to these large changes that really are out of your control for the most part. So I would say that in general the most difficult thing to do is to secure the long-term commitment of a certain amount of money forever” (RepositoryStaff\_09).

Repository staff members described a variety of political, organizational, and communication issues that posed threats to their long-term funding, including relationships with member organizations, staffing changes within their own institutions, competition with other repositories, and reliance on short-term funding sources such as grants. In the case of member funding, one repository staff member explained that staff turnover at member organizations meant that representatives from those institutions often misunderstood the relationship between the repository and their own institution, which threatened the repository’s funding as a result:

*“We're member funded. Our members, so we have members that have, the institutions have been members of us for a long time. Some of those institutions were founding members. But the individual people who are representatives for those institutions change fairly often, and one of the things that we're constantly dealing with is that we have to reeducate our members as to the fact that they're members. They forget that they're members and they think we're a vendor. So the biggest problem we have is that the people who fund is think we're a vendor. And so we have to constantly go back to them, including some of the very organizations that were involved in founding us in the first place, and remind them that, 'No, you founded us to solve this problem. No, we're not a content vendor. Don't just sit there and say- ' So we've had problems where we've had people that were on our board that were on the boards of other organizations that were negotiating with us to lower our fees. And it's like, why would you even negotiate to lower our fees when you should be negotiating to make sure that we have strong, long-term, stable funding?’” (RepositoryStaff\_16).*



Other repository staff members reported that uncertainty about future funding for their repositories were potential sources of risk for both the repository and the digital content that they were preserving. For Repository Staff 17, uncertainty about future funding was a threat, “I would say, the funding is the greatest threat ... we're quite uncertain of our future right now, and that's a bit stressful” (RepositoryStaff\_17). And for Repository Staff 19, uncertainty about where funding would come from at the end of each grant cycle was a potential source of risk, “That was probably the biggest one, that was coming up with something that's going to let this live beyond the next funding round” (RepositoryStaff\_19).

Repository Staff 04 spoke about his repository at the time of their TRAC audit and explained that they were entirely reliant on short-term grant funding. As a result, they received a low score from the auditors in the area pertaining to long-term financial sustainability:

*“It's always been the money is the hardest, most riskiest challenge. And that's one of the reasons we've pushed hard to make at least a core of what we do in [repository] part of our library organization ... Whereas previously, when we were getting audited, that came all from grant funds. Everything at that time had a very finite horizon on it. That is actually one of the lowest scoring category we got in the TRAC process, was because of our funding.” (RepositoryStaff\_04)*

He went on to explain that this area of the audit was difficult for his repository, because they understood the requirements laid out in the TRAC standard, and knew what types of documentation the auditors wanted to see, but were unable to meet those requirements in areas where his repository was vulnerable:

*“Difficult was the financial stuff, just because honestly, we didn't have a lot to say. It was difficult to answer the kind of questions that were being asked. Because that was where we were on the shakiest ground ... Because we were 100% grant funded and we knew the grant was going to end. And so we had a lot of answers of well, here's what we're going to try. But it was difficult to know okay, how do you actually say that without sounding stupid? Or without them saying well clearly, this is the way you do it. So that was really difficult to answer” (RepositoryStaff\_04).*

Repository Staff 20 and 21 also said that uncertain finances were a problem for their repositories at the time of their TRAC audits. For 21, uncertainty about funding sources meant that they were unable to demonstrate the long-term viability of the repository, “the far future part, who was ensuring that the process will keep going ... It was uncertain where the funding was going to come from, I guess, that was a big deal, the funding” (RepositoryStaff\_21). And for 20, reliance on their university meant that while their repository may have appeared to have a stable funding source, in reality the repository’s funding ran on two-year cycles, “our whole financial situation is kind of, you know, we have to come up with a budget every two years and justify all the salary and hardware and all the other expenses. And, ask the school to pay for it” (RepositoryStaff\_20). In both cases, the repository staff understood the risk posed by lack of long-term funding for their repositories, but were unable to meet the requirements described in the standard to mitigate that risk.

For repository staff, threats to the financial sustainability of their organizations were described as a major source of potential risk. They expressed concern about securing funding for their repositories, and frustration about their inability to meet the level of funding and long-term security that the TRAC standard described. While many explained that their repository fell far short of the requirements described in the standard, their repositories were still certified by TRAC auditors.

With regard to financial uncertainty as a potential source of risk for digital repositories, standard developers and auditors, groups composed of individuals with advanced degrees and technical expertise whose professional roles include administrative and management responsibilities, described the need for stable, long-term funding and acknowledged that securing this type of funding is difficult for digital repositories. Auditors agreed with the criteria

established by the standard developers that called for repositories to provide evidence of business planning processes and financial risk analysis reporting in order to demonstrate risk mitigation in this area. Repository staff, a group with greater variety in both their educational backgrounds and professional responsibilities, who tended to have fewer advanced degrees than the standard developers or auditors, also described financial uncertainty as a potential source of risk for digital repositories, but disagreed about whether the TRAC criteria could effectively assess the financial sustainability of a repository.

Uncertainty, vulnerability, and communication surfaced as prominent factors in the construction of financial risk for digital repositories. Standard developers, auditors, and repository staff all described funding as uncertain, and digital repositories as vulnerable to funding sources over which they had little control. Repository staff members, a group that consisted of individuals with a variety of roles and responsibilities within digital repositories but who mostly shared an educational background, also characterized their efforts to communicate with funders and parent organizations about the cost of long-term digital preservation as highly challenging.

While much of the focus on financial sustainability was about ensuring that a repository had the resources and plans in place to carry out its mission of long-term digital preservation, the TRAC standard also stated that a repository must have a succession plan in place in order to ensure the longevity of the digital content in the event that the repository was unable to do so.

#### **4.3.1.1 Succession Planning**

Checklist item 3.1.2.1 in the TRAC standard specifies that a repository must be able to provide credible evidence documenting a succession plan in the event that the repository loses funding and must shut down. While this requirement is included here in the financial risk

section, it could also fit with the sections about governance and legal sources of risk below. From a financial perspective, succession planning is important because a repository that loses funding will need a plan to ensure the longevity of its digital content, although repository staff members expressed skepticism that another repository would be able to secure funding if the first shuts down. Succession plans may also be necessary if a repository loses organizational support, or if the goals of the organization shift away from long-term preservation of digital content. And finally succession planning relies on legal agreements between organizations, although repository staff members interviewed were skeptical that those agreements would be enforceable.

The text of item 3.1.2.1 provides information about this item, including examples of how a repository can demonstrate that it has met the requirement:

*“3.1.2.1 The repository shall have an appropriate succession plan, contingency plans, and/or escrow arrangements in place in case the repository ceases to operate or the governing or funding institution substantially changes its scope.*

#### *Supporting Text*

*This is necessary in order to preserve the information content entrusted to the repository by handing it on to another custodian in the case that the repository ceases to operate.*

#### *Examples of ways the Repository Can Demonstrate It Is Meeting This Requirement*

*Written and credible succession and contingency plan(s); explicit and specific statement documenting the intent to ensure continuity of the repository, and the steps taken and to be taken to ensure continuity; escrow of critical code, software, and metadata sufficient to enable reconstitution of the repository and its content in the event of repository failure; escrow and/or reserve funds set aside for contingencies; explicit agreements with successor organizations documenting the measures to be taken to ensure the complete and formal transfer of responsibility for the repository’s digital content and related assets, and granting the requisite rights necessary to ensure continuity of the content and repository services.*

#### *Discussion*

*A repository’s failure threatens the long-term sustainability of a repository’s information content. It is not sufficient for the repository to have an informal plan*

*or policy regarding where its data goes should a failure occur. A formal plan with identified procedures needs to be in place.” (Consultative Committee for Space Data Systems, 2012, p. 22)*

This checklist item provides further evidence that the developers of the TRAC standard identified financial sustainability as a potential source of risk for digital repositories, and that they agreed that a credible succession or contingency plan would serve as sufficient evidence of a repository’s ability to respond to the risk of financial instability. Interviews with standard developers confirmed the view that documented evidence of plans to preserve and provide access to digital content beyond the life of the repository was a critical strategy for mitigating financial threats. For example, Standard Developer 09 found financial viability to be a potential source of risk for digital repositories and explained that a succession plan would be needed to mitigate this threat:

*“A good repository, even if it's making assumptions about its user base, even if it's gone a bit complacent, it would probably notice eventually that things are changing and take some steps to act on that. Whereas, risks concerning viability — I'm thinking of financial viability — can come out of the blue. I mean, there have been instances of repositories simply being closed ... I would say that there is a case to be made that that's the greatest risk, because it's unpredictable, and it could wipe out a whole repository as a stroke, which is why, of course, you need succession plans.” (StandardDeveloper\_09)*

Despite the emphasis on financial sustainability described in the previous section, standard developers also recognized that securing long-term funding was a significant challenge for repositories. Standard Developer 07 explained that the requirement that a repository have a succession plan was meant to provide an additional or alternative way for repositories to demonstrate the longevity of their digital content, “All of those sorts of things, and other repositories, the difficulty is the long-term funding, so in OAIS, the 16363, we kind of get around that by talking about having a succession plan” (StandardDeveloper\_07).

In addition to having a succession plan, standard developers argued that the plan should be viable and actionable, and that the organization designated to take over the digital content should be able to do so. For example, Standard Developer 05 explained that a succession plan should be both well-developed and thoroughly documented in order to ensure that it will be actionable in the event that it is needed, “Disaster recovery and the succession plan must be well documented and, really, very concrete in ensuring the continuity and the persistency. In case of failure, in case if the repository closes, in case the company who has the storage system decided not to do it anymore and disappear. It happens. It can happen. So, the succession plan must be really well, well, well, well developed” (StandardDeveloper\_05). She went on to say that a succession plan encompasses the other risks that a repository faces, and that a solid succession plan is evidence that a repository has addressed other major categories of risk, “The succession plan, in some sense, include[s] all the other risks. If you don't have a visible, credible, trustable succession plan, you cannot really provide a good – you cannot face your challenge. But the succession plan means that all the other things must be there, in some sense” (StandardDeveloper\_05).

Standard Developer 10 also focused on the viability of succession plans, and argued that the repository in the vignette would need to provide sufficient evidence that the organization designated in the succession plan would be capable of taking on the new content, “We’d have to know what is the viability of that succession plan, what other place is capable of taking this on” (StandardDeveloper\_10).

As with the standard developers, auditors described succession planning as an important and necessary measure for repositories to take in order to offset the potential risk of organizational collapse and loss of digital content, “I think that, in terms of the organization, they

need to develop a succession plan and be very explicit about what's going to happen if their grant funding dries up, and if the membership starts to drop” (Auditor\_10).

Auditors tended to frame their discussion of succession planning in terms of their audit experience rather than their digital preservation or repository management experience. They described succession plans as necessary for long-term digital preservation, but also said that it was uncommon to find repositories with excellent succession plans, “Well first of all these days, having a succession plan is really good ... I mean it’s pretty rare, actually, for a repository to actually have a succession plan. So the fact that they have one is good news” (Auditor\_06).

Auditor 08 explained that in his experience as an auditor the existence of a viable succession plan was a good indicator of the overall state of the repository. This is similar to the point made by Standard Developer 05 above, who argued that a succession plan was in some ways an indicator of the repository’s overall preparedness to address potential threats. For Auditor 08, whose understanding of financial risk and succession planning was shaped by his experiences assessing repositories, it was important to see that repository staff recognized the potential sources of risk that could threaten the financial sustainability of the organization and had made plans for the digital content to ensure that it would outlive the repository itself, “I don't think organizations always think of succession planning when they're doing this. From your point of view, as say an Executive Director of an organization, you believe in the vitality and health of that organization and that it will persist into the future, and doing that planning upfront of what happens if we decide to dissolve or we’re no longer organizationally viable, whether it's money or personnel or something else? The very best audits thought of those issues, and the ones that were most problematic, that wasn't on their radar screen at all” (Auditor\_08).

Taking that a step further, Auditor 01 said that while it was important to know that the repository had a succession plan in place, it was also necessary for the repository to have tested that plan to ensure that transfer of digital content to the new organization would be possible, “Then the succession plan that they say that they have. Has that been tested? How many times have they tested that? What kind of variety of data have they tested it with?” (Auditor\_01).

As with the standard developers, auditors described succession plans as necessary tools to offset threats to the financial sustainability of a repository. While standard developers emphasized the importance of well-documented succession plans, auditors noted how difficult and rare it was for a repository to have such a plan. Auditors also highlighted the difference between having a documented succession plan and carrying out testing to ensure that it would be possible to transfer content to the new repository if the need arose.

Repository staff agreed with the standard developers and auditors that financial sustainability was a potential source of risk for digital repositories and the content that they sought to preserve, “There's always a risk in that, with whatever might happen to that organization. Either a calamity, or loss of interest, or will, or funding, or whatever. There is a succession plan it says in there, so that's obviously a significant mitigating tool for that kind of failure of the organization. I think succession is tricky” (RepositoryStaff\_06).

Echoing the sentiments of Auditors 06 and 08, Repository Staff 05 said that while funding challenges are a common and substantial threat to digital repositories, in his experience most repositories do not have a succession plan, “I think a lot of institutions have been facing significant funding challenges ... if the funding dries up, then support for that repository, you know, what happens? You've got that data. Do you even have a succession plan? I think a lot of places don't” (RepositoryStaff\_05).



Repository staff disagreed with standard developers and auditors about whether a succession plan was sufficient evidence of risk mitigation. Repository Staff 03, 06, 07, 12, and 21 all expressed skepticism that having a documented succession plan would ensure the longevity of a repository's digital content, "I wasn't necessarily convinced that writing that down necessarily meant that it would sustain it" (RepositoryStaff\_03).

Repository Staff 03 said that his repository was not concerned about succession planning at the time of their TRAC audit because the repository was covered by the university library within which it was situated, but that as the repository has grown and become more independent succession planning has become more complicated and nuanced, "We were never really that concerned about succession with major universities. But, I think, this is the case with [repository] as well, how that staff entity sits within the university environment has changed due to organizational, political realities. Whereas, when we were incubating the whole service, it was very tightly tied, almost indistinguishable from library IT, and now it's really starting to become more of its own entity. I think that's more of the nature of the succession plan. It's not so much, 'Oh my gosh, what happens if the [university] suddenly goes under?' It's a little bit more nuanced" (RepositoryStaff\_03).

Similarly, Repository Staff 21 said that her repository did not have a succession plan at the time of their TRAC audit. She explained that she thought they should have one because it was included in the checklist and she knew that the auditors would expect it, but that she doubted whether a succession plan would meaningfully address the potential risk of financial instability, "we didn't really have a succession plan. I don't know if they do now. I'm not quite sure. But at the time we didn't have one, and I remember thinking, oh, we probably should have that, even if it's just written up. I don't know how meaningful it would have been if we just threw one

together, but that was a big deal, because if we lost funding, then what was going to happen to all that content?” (RepositoryStaff\_21).

Repository Staff 12 was quite blunt in her assessment of succession planning as a futile activity. She argued that succession planning did not make sense because it is unlikely that a second repository would be able to muster the funding and support that the first was lacking, “What is really going to be the reason repositories are at risk, is almost all around having enough money to take care of the material . . . a succession plan to move it someplace else, where the community isn’t going to have enough money to take care of it. Or there’s going to be a, someone who magically dumps money on the secondary repository. Why couldn’t they dump money on the first repository? I mean, it’s just, I don’t know. It doesn’t make sense” (RepositoryStaff\_12).

Repository Staff 07 went a step further and explained that by their very nature succession plans are unenforceable because they are only enacted when a repository fails. He also said that a succession plan does not ensure that the successor organization itself will be financially viable long-term:

*“It was always about sustainability and succession. It was basically having a viable plan for what would happen to the content in the event that we were unable to continue operating as an organization and had to wind up, or just couldn’t support the repository anymore. Having basically succession partners who would not only take the content and the infrastructure around it but also operate it. And I think that’s still sort of, it’s almost like that’s a weak link too because if you have a succession by definition you’re gone afterwards so you can put a plan in place but you’re not around to make sure that it’s going to be executed. Just like you’re not around forever your successors aren’t necessarily around forever. Our successors are primarily universities and government agencies which all claim and pretend that they will exist forever, but you can’t guarantee that so the succession plan doesn’t actually spell out what’s going to happen from now until the end of forever it just says that there’s an agreement in place, it’s a time limited agreement. The idea behind it would be that our successors would agree that they would operate our infrastructure and maintain access to the content, and preserve it while the next steps were investigated. [repository] itself, as an organization,*

*we don't even have the ability to pretend we're going to be around forever because our funding base is much smaller and more fragile than these other organizations” (RepositoryStaff\_07).*

Standard developers, auditors, and repository staff all agreed that loss of funding and/or institutional support was a potential source of risk for digital repositories and their content. Interviews with all three groups revealed similar ideas about the likelihood and consequences that a repository could lose funding. Standard developers and auditors, two groups whose members were likely to occupy professional roles with policy setting responsibilities, emphasized the importance of succession plans as evidence that a repository was prepared to address this potential source of risk by ensuring the longevity of digital content beyond the life of the repository itself. Repository staff on the other hand, a group consisting of members who were likely to occupy professional roles in which they would be responsible for carrying out the work specified by policies that they did not create themselves, understood the reasoning behind succession planning but did not agree that a succession plan provided evidence that the digital content would outlive the repository. While they were happy to provide documented succession plans in order to achieve TRAC certification, they were performing rather than demonstrating trustworthiness.

#### **4.3.2 Legal**

Interviewees described legal issues, such as contracts, agreements, licenses, and copyright as potential sources of risk for digital repositories. While auditors and repository staff members agreed with the conceptualization of legal risk presented by the standard developers, the auditors were more likely to agree with the standard developers that agreements among organizations governing relationships that would impact the long-term preservation of digital content should be the primary focus of concern. Repository staff members were much more

concerned with intellectual property issues that would threaten the repository itself. It is understandable, in this case, that repository staff would be concerned with the ability of their own organization to do this work, while individuals external to their organization would be more interested in the relationships with other organizations. This echoes the attitudes expressed in section 4.3.1.1 about succession planning. Standard developers and auditors believed that it would be possible for digital content to outlive the repository, while repository staff were skeptical that succession plans would be successful.

Standard developers framed legal risk to repositories as something that was of particular importance to repositories in relation to access. For example, Standard Developer 01 explained that the legal repercussions of releasing protected data could threaten the continued existence of a repository, “There’s laws in place in the U.S. I don’t know about the rest of the world, but certainly in the U.S., depending on what your repository is storing you may have very severe penalties imposed on you if you release information that’s supposed to be protected. The HIPAA [Health Insurance Portability and Accountability Act] is one example. There’s a Title XIII, which is census data. Both of those are legal system where keeping the data under security controls is tantamount to keeping your organization from being ground by the wheels of justice.”

Alternately, Standard Developer 07 said that the team of standard developers were not concerned with threats posed to a repository by legal issues, but rather to the digital information being preserved by that repository, “The problem there is that if the repository could be sued, legal action taken against them, then we had a very shorthand way of describing this: that we didn’t care if the repository itself was sued out of existence. What we were concerned about is that they were sued out of existence before it could hand over its data, its information. It’s the speed at which these legal activities, cease and desist operations, can happen.” For this

interviewee, the danger to the organization, that a repository would be shut down before they could enact their succession plan, was a significant threat to the digital content.

In order to guard against the legal risks that repositories face, standard developers argued that it was important to have clear, well-defined, agreements, “From your Organizational Infrastructure an issue is to make clear which kind of agreements you have within the organizations involved, other centers they could be in contact with to preserve things, to receive things. And, some specific agreement must be done on a very clear basis. The duties, the licenses, the risks, the copyright. All these things must be, again, well defined . . . Evidence is very important in this area” (RepositoryStaff\_05).

This view of legal risk is clearly described in section 3.5 of the TRAC standard, “Contracts, Licenses, and Liabilities,” which states that “The repository shall have and maintain appropriate contracts or deposit agreements for digital materials that it manages, preserves, and/or to which it provides access” (Consultative Committee for Space Data Systems, 2012, p. 31). Through the process of creating this text the standard developers established an understanding of legal risk as one that was a threat to both the repository and the digital content, and communicated to both auditors and repository staff members that it was necessary and important to “ensure that the repository has the rights and authorizations needed to enable it to collect and preserve digital content over time, make that information available to its Designated Community, and defend those rights when challenged” (Consultative Committee for Space Data Systems, 2012, p. 31).

Standard developers set expectations for auditors and repository staff that a repository could demonstrate that it is meeting this standard by providing documentation, “Properly signed and executed deposit agreements and licenses in accordance with local, national, and

international laws and regulations; policies on third-party deposit arrangements; definitions of service levels and permitted uses; repository policies on the treatment of ‘orphan works’ and copyright dispute resolution; reports of independent risk assessments of these policies; procedures for regularly reviewing and maintaining agreements, contracts, and licenses” (Consultative Committee for Space Data Systems, 2012, p. 31).

Through the text of the TRAC standard, the standard developers crafted and communicated a conceptualization of risk with regard to legal issues that focused on the longevity of digital content rather than the security of the repository. They were concerned that repositories would obtain the appropriate permissions to be able to preserve digital content, and that they would have secure enough legal footing to ensure that they would be able to execute their succession plan in the event that the repository should fail.

Auditors did not devote much attention to legal risk during the interviews, but their discussions about legal issues tended to express a shared understanding of risk in this area with the standard developers. Drawing from their experiences conducting audits as well as their own professional backgrounds in digital preservation, they focused on one aspect of the legal risk communicated through the TRAC standard. Namely, that it was important for repositories to have the appropriate legal agreements in place in order to ensure that their relationships with partners and members were secure.

Auditor 01 said that repositories “should probably have some legal staff on hand” to manage contracts among partner organizations because negotiating and executing things like service level agreements is complex and time-consuming. He also said that when assessing a repository it is important to understand whether those agreements are reciprocal or not in order to fully understand relationships among organizations and the potential sources of legal risk that the

repository faces, “Is this a reciprocal agreement and what kind of risks does that expose them to?”

Auditors also framed legal risk in terms of the position of the repository with regard the hosting of its digital content and support for the mission of long-term preservation, “Just the word support is not enough. There needs to be some formality around that and documentation agreements” (Auditor\_09). They said that it was important to understand how the digital content would be preserved in the event that the organization hosting that content decided to withdraw from the repository, “What if the founding university decided to withdraw from the project? Then they wouldn’t want to host the content anymore, so you need to look into whether or not there’s actually any contracts specifically about the university agreeing to host the repository infrastructure and content and what they say in the event of there being a disruption or if they withdraw” (Auditor\_03).

While standard developers identified legal risks to digital repositories arising from copyright issues as significant, they explained that it was necessary for repositories to secure the rights necessary for long-term preservation of the digital content. Repository staff, however, were more concerned that even if these legal agreements were in place, execution of the access permissions and/or restrictions specified in those agreements would somehow fail, “a lot of the complexity came from ... being able to provide access in the right ways” (RepositoryStaff\_01).

Repository staff presented a view of legal risk that included a great deal of concern about copyright and the threat posed to repositories that provided inappropriate access to digital content. Repository Staff 06 explained that “the risk of compromise to the content that's in copyright” was an area of vulnerability for repositories. For this interviewee, the threat of providing inappropriate access to materials with copyright restrictions was a potential legal threat

to a repository. He went on to argue that access in general is an area of risk for repositories, and that the push to provide repository users with meaningful ways to access and interact with data can interfere with the core mission of preservation by pulling resources away from that work, “I think access in general is complicated and getting more complicated.” Extending that perspective, Repository Staff 01 said that an important question for repositories was, “What is your capacity to actually provide someone something they can use within the constructs of law that exists?”

Intellectual property rights were described as a “ticking time bomb” by Repository Staff 02, who explained that repository cost models were complex sources of potential risk for repositories, “The way that national copyright factors into the cost model, which is two-dimensional and I think very complicated, but it probably needs to be multidimensional more than that because of copyright issues.” Despite this concern, he felt that the auditors who assessed his repository had an inflated sense of the threat that copyright issues posed to his repository. He said that he disagreed with their “sense of risk” with regard to in-copyright materials, but “didn’t feel it was worthy of dispute” in the final TRAC audit report.

Digital content with access restrictions due to copyright were described as a source of vulnerability for repositories because if a repository was hacked it could “lead to lawsuits and things like that where that could really jeopardize the stability of a repository”

(RepositoryStaff\_05). This interviewee argued that, “You can have a great technological plan that can quickly become irrelevant if you haven’t got all of the legal ducks in a row”

(RepositoryStaff\_05). In contrast, Repository Staff 13 said that technical aspects of repository architecture could eliminate the need for complex legal agreements regarding intellectual property rights because “the fundamental problem we were facing was copyright” and “by



throwing large amounts of disk at the problem, we could avoid throwing large amounts of money at lawyers.”

As with the standard developers and auditors described above, a few repository staff members were also concerned with agreements governing relationships between repositories and partner or member organizations. They described these relationships, and the agreements governing them, as complex:

*“One area of complexity is in the legal realm, and in the agreement realm. The complexity there is the different regimes that you have to deal with between say, private universities and public universities. The different laws of different states. The [state] is a complexity. That really is complex, and how to try to simplify that into an agreement that someone can actually sign, and that can be uniform across your services. That's one degree of complexity. The other, I think it also stems from the legal things in a way, is what [repository] was trying to do was preserve and provide access. They're very interconnected.” (RepositoryStaff\_01)*

Repository Staff 18 explained that the agreements governing his repository were complex and required a great deal of time to understand, “The way that [repository] operates is through a service provider agreement that seems a bit Byzantine when you first start looking at it ... So there's a level of service agreement there that's complicated and kind of hard to get your head around until you've been in it for a while.” One implication that follows from this description is that it could be difficult for an external auditor to understand the legal landscape of this repository and to, in turn, assess risk in this area.

Discussing the TRAC audit process specifically, Repository Staff 19 highlighted the importance of organizational factors and argued that an audit would need to look across all of the partner/member organizations hosting digital content for a repository, “I would have to say, you'd have to put in some agreement as how do you co-audit across them.” Regarding his own repository, he explained that the audit team spoke with staff at all of the sites hosting content and reviewed documentation but only visited the primary site, “I don't believe they ever visited the

other [number of sites] sites.” Another interviewee from this repository said that he was under the impression that auditors had visited at least one of the other sites hosting their content during the audit process. In the case of this repository, staff members within the same organization had different understandings about whether the auditors had visited partner organizations or not.

Several interviewees identified national context as a potential source of risk for the digital repository described in the vignette. Of the two standard developers, three auditors, and five repository staff members who thought that having a data backup location in a different country was a potential source of risk, five were from the U.S., and five were from Canada and Europe. Standard developers, auditors, and repository staff agreed that having data storage sites in two different countries could endanger or limit federal and/or state funding for repositories, could introduce complexity by adding additional laws and/or regulations that a repository must comply with, and could create problems for material under copyright.

While standard developers, auditors, and repository staff all found legal issues, such as contracts, agreements, licenses, and copyright, to be potential sources of risk for digital repositories, the groups focused on different sources of risk. Standard developers and auditors focused on relationships among partner and/or member organizations, and argued that those relationships were a potential threat to both repositories and digital content, and that agreements were necessary in order to ensure and enforce a commitment to the mission of long-term digital preservation. Repository staff, on the other hand, focused primarily on intellectual property issues and the threat that violating copyright posed to their repositories. They also spoke about the complexity of the legal agreements governing relationships among partner and/or member institutions and expressed some skepticism about whether an external party would be able to understand the legal landscape of their repositories.

Although all three groups involved in the TRAC audit process agreed that legal issues were important, the fact that repository staff members focused on a different aspect than the standard developers and auditors suggests that individuals who are removed from the process of repository management find potential risk in the instability and uncertainty of relationships among member and/or partner organizations while the individuals who are in repositories are more concerned about potential threats from people and organizations external to the repository and its partners. Repository staff were focused on TRAC certification as a marker of whether a specific repository could be considered a trustworthy home for digital content, while standard developers and auditors focused on certification as a marker of how likely it was that digital content could outlive the repository itself.

#### **4.3.3 Organizational Governance**

Interviewees described organizational instability as a potential source of risk for digital repositories. They discussed the ways that internal organizational governance structures and the positioning of the repository with larger organizations, such as universities, consortia, or partnerships as possible threats to both a repository and its digital content. While standard developers focused on the ways that the requirements laid out in the TRAC standard would mitigate potential threats relating to organizational instability, auditors and repository staff were skeptical about whether policies and documentation could accurately capture actual repository practices. Auditors emphasized that a focus long-term digital preservation, and a mission statement and policies that reinforced this focus, would mitigate the potential risk of organizational instability. However, repository staff described TRAC certified organizations without clear mission statements, and where staff members lacked a clear understanding of the overall mission of long-term preservation.

Section 3 of the TRAC standard focuses on organizational infrastructure and includes several sub-sections that specifically target governance, including section 3.1 “Governance and Organizational Viability” and section 3.2 “Organizational Structure and Staffing.” The Governance and Organizational Viability section specifies that a trustworthy repository should have a mission statement that reflects a commitment to digital preservation, as well as a strategic plan, a succession plan, and a collection policy that all reflect the mission of long-term preservation. The Organizational Structure and Staffing section focuses on the need for appropriate staffing to carry out the mission of long-term preservation, “3.2.1 The repository shall have identified and established the duties that it needs to perform and shall have appointed staff with adequate skills and experience to fulfill these duties” (Consultative Committee for Space Data Systems, 2012, p. 23). The standard indicates that in addition to having appropriate staff, a repository should have clear documentation communicating the duties that staff should perform, evidence that there are enough staff members to carry out the work, and evidence of ongoing training and professional development in areas relevant to digital preservation.

The requirement that a repository provide documented evidence of a mission focused on long-term preservation of digital content, along with supporting documentation describing policies that align with that mission reflects a belief on the part of the standard developers that loss of institutional support for digital preservation and/or shifting organizational priorities are a threat digital repositories, “The repository’s or its parent organization’s mission statement should explicitly address preservation. If preservation is not among the primary purposes of an organization that houses a digital repository then preservation may not be essential to the organization’s mission” (Consultative Committee for Space Data Systems, 2012, p. 21).

Standard developers identified three areas of organizational governance as potential sources of risk for digital repositories: (1) institutional support, (2) leadership changes, and (3) organizational structure. Loss of institutional support was described by several standard developers as a major threat to digital repositories. Standard Developers 01, 02, 05, 06, 08, and 10 all emphasized the potential risk for repositories and digital content associated with loss of support for the mission of long-term digital preservation.

When discussing the repository described in the vignette, Standard Developer 02 explained that turnover among board members was a potential source of risk for the repository because board members develop knowledge and expertise about the organization over time, new members may not understand the ways that their decisions influence the mission of long-term digital preservation, “Because the board is possibly all new members at any given time, they're not going to necessarily see so much of the change that's taking place and might not even, even if they're told about it, they might not understand the implications in terms of the long-term objectives for preservation and access. To me, that could be a little risky” (StandardDeveloper\_02).

Standard Developer 05 said that uncertainty about organizational structure and staffing was a potential source of risk for digital repositories, “I think that the main question of uncertainty is related to the low level of organizational infrastructure, more than any other thing. Because if you have good people, at the right point, and the responsibility is well developed, the uncertainty could be covered” (StandardDeveloper\_05). This attitude toward organizational infrastructure and the emphasis on appropriate staffing of people with expertise that is relevant to the repository reflects the requirements described in the TRAC standard, which emphasize the

need for clear policies governing the structure of the repository and ensuring that staff understand how their roles support the overall mission of long-term digital preservation.

This view of organizational infrastructure and governance as a potential source of risk for digital repositories reflects a view of digital repositories as organizations that are at risk of losing focus on long-term digital preservation either because of mission scope creep or because parent or partner organizations may have goals that differ from the repository. In this sense, the standard sets an expectation that repositories will need to defend their focus on long-term preservation and that repository staff members should all understand how their role serves that mission.

TRAC auditors reinforced this conceptualization of governance as a potential source of risk for digital repositories, focusing on institutional support as the most prominent potential source of risk, “the most important aspect of a repository is having an organizational commitment with a mission that aligns with the repository” (Auditor\_06). Auditors 01, 03, 04, 05, 06, 07, 08, and 10 described governance and organizational stability issues as both complex and uncertain:

*“We don't know if libraries are going to survive. We don't know if universities are going to survive. These institutions that support the institution, the repository, are also at risk at this point. There's no certainty of anything surviving, so requiring that particular content survive into the future, it's really going to be hard to figure it out, where to go and who should be supporting this kind of work. Yeah, I do think that we don't really know where things are going. We've constructed this organizational structure that includes digital repositories, but the content itself, whether or not it's protected or not, I don't really know where ... I don't know who's going to support it in 50 years. I don't know if it's still going to be a library or a university or it's going to be some crowd funded thing. I just don't know, so I think that is the biggest risk for almost everything that we're doing now is knowing what's going to happen to these institutions because a lot of things are at risk right now.” (Auditor\_03)*

Auditors expressed attitudes similar to the standard developers when discussing the importance of governance in a TRAC audit. Reflecting the requirement described in the standard

that digital repositories should have explicit mission statements emphasizing long-term preservation, for example, auditors described ongoing organizational support for preservation as a challenge for repositories, “Bottom line is it's a tremendous amount of resources required to do long-term preservation. Organizational commitment to those types of resources often waxes and wanes” (Auditor\_05). This auditor went on to say that he thought that the organizational infrastructure elements of the TRAC checklist were more aspirational than realistic because in practice repositories lack support for long-term preservation:

*“I think certainly the organizational stuff is probably the ... it's the most aspirational simply because there are very few organizations who are anywhere near in that place that they've got the commitment and the wherewithal and the long-term view. In that sense, I'd say the organizational side of it. The reality is that there are, for instance, there are very few organizations who have something like a formalized digital preservation unit. You have very few organizations that have formalized procedures and plans and have those things incorporated into let's say their operating goals. That actually is probably addressing a very, very small number of organizations who are positioned and capable of getting to that place. And typically they're very large. They're your national libraries, they're your large collectives. They're at that level, so you're talking about a very small number of organizations” (Auditor\_05).*

This auditor highlighted the difference between the ideal described in the TRAC standard and the reality for digital repositories. The standard states that a trustworthy digital repository should have a mission focused primarily on long-term preservation with policies and a staffing model designed to support this mission and institutional backing by parent and/or partner organizations that share this mission. In practice, auditors described repositories as organizations with competing priorities who must continually fight for resources to support long-term digital preservation efforts, and whose parent and partner organizations may or may not share their commitment to preservation.

A TRAC audit reflects a snapshot of a digital repository at one point in time. Auditor 07 explained that securing a commitment to organizational support is an important step, but

maintaining that support over time is challenging, “I think things are changing all the time, so it's really having the organizational will and money and time to put towards maintaining it. Getting it is the first step, but maintaining it is the critical ongoing work.” This view challenges the notion that a one-time audit can assess whether a repository is capable of maintaining organizational support for long-term digital preservation.

Other auditors expressed concerns about repository governance that reflected a similar attitude, describing organizational support as changeable, “So I guess what I'm trying to illustrate is that bad times can come in clusters. It's not just one university saying we can't afford to join this year. It could be a bunch, depending on contextual or jurisdictional circumstances. That's a major risk for any, being a dues-based membership system organization is a risk that needs to, it's not an insurmountable risk, but it's one that I think an organization would need to be aware of and would need to disclose in its certification process” (Auditor\_09).

In light of the changeability of organizational support, Auditor 03 said that the audit team would want to be alerted to major changes in repository governance over time, “if they were to change their governance or if some big change happened, then we would definitely want to know so we could discuss it.” This fits with the approach to certification outlined by the auditors – namely, that repository staff members are responsible for contacting CRL to notify them about major changes to the repository in order to maintain TRAC certification over time. However, standard developers and repository staff did not have this expectation. Rather, both standard developers and repository staff expected certification to include periodic audits that would be conducted on a schedule set by the auditors. Indeed, most repository staff members who were interviewed said that they were expecting to be contacted by CRL for a recertification audit at some point in the future.



Overall auditors shared the standard developers' view of organizational instability as a potential source of risk for digital repositories. While the standard developers described, through interviews as well as in the text of the TRAC standard itself, strategies for a repository staff to demonstrate that they had policies and procedures in place to mitigate this risk, the auditors took a more circumspect approach to verifying that repositories were mitigating this potential source of risk. They described institutional support for digital repositories as changeable and likely to decrease over time, and explained that it was easier for repositories to secure initial support for digital preservation than to maintain support over time. Auditor attitudes about governance as a potential source of risk questions the notion that a one-time audit can assess whether a repository should be considered trustworthy in its ability to preserve digital content over the long-term. Auditors are enforcing requirements from the TRAC standard in order to certify a repository as trustworthy, but they are skeptical about whether long-term trustworthiness with regard to governance can be determined in this way.

As with the standard developers and auditors, repository staff members described governance and organizational stability as potential sources of risk for digital repositories, "I feel like the funding, the organizational governance, all those things are inherently risky and problematic" (RepositoryStaff\_02). Like the auditors, these interviewees questioned whether TRAC certification could assess the stability of repository governance over time, "I think it probably could be quite difficult for any kind of certification program to validate how functional a governance system is" (RepositoryStaff\_05). Repository staff members described policies and practices at their organizations that were complex and continually evolving.

While all of the repository staff members described long-term preservation of digital content as important for their organizations, there was disagreement about whether this should be

the central mission of the repository. One interviewee in particular reported that his repository did not have a mission statement, and that their long-term goals focused on meeting user needs, which happened to include providing long-term access to particular content that was of interest to their Designated Community. In the documentation that his repository provided to auditors, the goals of their preservation efforts were articulated in the description of their Designated Community as providing long-term access to specific digital content for that community, but these preservation efforts were not described as part of the repository's mission. Repository Staff 18 described the workaround that his repository used to address the criteria in the standard without creating a mission statement for the repository that focused specifically on long-term preservation:

*“One thing that was interestingly difficult to get was a sort of mission vision statement. Which is on some level kind of the most basic of the documents that can be. Because it established why you exist in the first place. But it turns out we and a lot of other organizations don't have that existing in that form. Rather our mandate and our vision comes out of ... well, mandate comes out of the fact that the schools continue to pay money to us to exist. And our vision comes from our governance structure. So on some level you can say that our vision is to do what our community needs us to do. But that's not really useful in the context of the audit, so figuring out a way to answer those questions with our strategic plan, which we do have, took some time and some conversation” (RepositoryStaff\_18).*

This runs counter to the position of the standard developers and auditors, who argued that a mission emphasizing long-term preservation was important and necessary for sustainable governance and organizational viability, and questioned the central premise of TRAC certification by asserting that a repository need not have a mission statement reflecting a commitment to long-term preservation of digital content.

Several interviewees, including Repository Staff 02, 03, 05, 14, and 15 discussed institutional support among partner and/or member organizations as a potential source of risk for digital repositories. Repository Staff 02, 06, and 14 explained that institutional support among

partners and/or members was a potential threat to the stability of a repository. Repository Staff 14 focused on political issues such as power struggles within the repository that could arise among members with a high level of investment in the organization, “There's a lot of negotiation that has to happen and I've seen power struggles with this big organizational infrastructure that I don't see on the technical side. Personalities get involved. I think that is the hardest part.” Alternately, Repository Staff 02 described problems that could arise when member institutions lose their sense of connection to the repository, “the institutional interest are always diffused through a governance process, through a prioritization process, through management. All of these things can result in the effort feeling less connected to the institution. The institution feeling less served by the effort.” Repository Staff 06 explained that the result of this lack of connection to the repository could be that participating members lose interest in the repository, “There's always the risk of the loss of sort of will or interest of the participating members, or at least enough of them. That could be expressed through disinterest in governing it” (RepositoryStaff\_06).

These examples highlight the balance of power that repositories must achieve in maintaining the interest and support of their partners and members in the overall mission of long-term digital preservation, without having any one organization or person take over. Maintaining institutional support from founding members was also described as a potential threat to repository stability in relation to repository growth over time, “I guess having a governance structure that is, we are kind of grappling with this point: How do you take a governance structure that has original members and then try and expand that and incorporate new members? Is there a tension between those original members and the type of control that they wanted to have over it, and the sort of deep engagement that I would imagine that they had as the

repository got started, and people who are probably likely to see it more as a service that they are paying for?” (RepositoryStaff\_15). One interviewee explained that having a large membership was one way for a repository to protect itself from the potential threat of losing members, but that this approach also had a significant drawback because members would not feel a strong sense of responsibility for the repository:

*“I think actually having a large number of organizations involved in paying membership dues provides a certain layer of security because you have the ability to survive the loss of any one or two particular organizations, but it also possibly represents a risk that with such diffused funding there's also a certain diffused sense of responsibility. And so if any of those institutions do come under financial strain they may feel that the repository can continue to exist without their contribution, or with a reduced contribution. Spreading out a funding base reduces the importance and impact of any given institution but it also, I think, decreases the potential sense of responsibility of each institution towards actually continuing to sustain and maintain that commitment.” (RepositoryStaff\_07)*

These interviewees depicted institutional support as a potential source of risk to repository sustainability because support from partner and/or member organizations was viewed as highly changeable and requiring ongoing attention, a position that echoes the attitude expressed by TRAC auditors. While institutional support at the time of a TRAC audit may be stable, the views described by repository staff members indicate that a one-time assessment of a repository will not provide a reliable appraisal of institutional support over time. For example, one repository staff member described significant changes to the composition of repository staff following their TRAC audit, “That audit was happening, too, during a project where we doubled the size of our staff. Then that project’s over, that has shrunk back down, so there are a lot of people who joined mid-audit and left either just after or just before the final certification too. Just actually thinking back to who was involved there were a lot of people who aren't here now who were definitely involved in it” (RepositoryStaff\_07).

Repository staff described staffing as a potential source of risk for digital repositories, and several interviewees either described their repositories as lacking appropriate staffing to carry out the work of long-term digital preservation, or described their fellow staff members as having a narrow view of the repository based on their roles rather than understanding their work in the context of the larger mission of long-term preservation. For example, Repository Staff 05 described challenges in concentrating IT staff within a repository on the mission of long-term preservation and said that his team has tended to lose focus on the larger mission of digital preservation, instead focusing on the shorter-term goal of software development, “Sometimes there's an argument that arises at times, that we're a software development shop. And, no, we're building a library. And we happen to need to make software to do it, right? And that's a hard concept sometimes for people to understand. I think their point is that we need everything that a software development shop needs. Yes that's true, but let's remember why we're here” (RepositoryStaff\_05).

Similarly, Repository Staff 21 described the staffing model of her repository as one that was too small to carry out the work necessary for long-term preservation, “It was an issue of having not enough staff.” She went on to say that most of her team members did not understand the concept of risk in the context of long-term preservation, “I don't think people there really, the people on the team aside from a few individuals, really understood that kind of concept of long-term risk outside of just technical stuff” (RepositoryStaff\_21).

Repository Staff 17 talked about staffing issues, explaining that staffing, policy, and transparency were all driven by available resources, and were therefore notably complex for her repository. She described her repository's TRAC audit as having started their process for creating appropriate staffing policies to support the work of the repository, but said that they

lacked the financial resources to have appropriate staffing, and to provide necessary training for the staff that they did have:

*“I'd say it's staffing, and sort of the roles within the TDR framework. We've had like a lot of cutbacks in the last year. I guess again this all relates to funding, but I think having the right people working in the right areas and the training that's necessary to sort of be really on the ball with all of these areas. I'd say that's a challenge. I think that effects, that's like intertwined in a lot of areas like to be really transparent in having policies in place for everything is an area I still think that really is complicated for us. I think the TRAC certification really helped us start that, like have a foundation, but I still find it a pretty big area that I think we could be doing better but we just don't have the resources.” (B17)*

Repository staff members described organizational instability as a potential source of risk for digital repositories, but were skeptical about whether documentation, such as mission statements could mitigate this threat. They were also skeptical about whether a one-time audit could accurately assess governance issues, which they described as highly changeable over time. While repository staff understood the TRAC standard's requirements in this area, and provided documentation to the auditors in response to the checklist items, they also described their repositories as lacking in actual governance structures, mission statements, strategic plans, and appropriate staffing. For example, although the TRAC standard called for clearly documented policies for organizational governance, Repository Staff 03 said that at the time of their TRAC audit, his repository had not yet established clear governance processes to manage relationships among member organizations, “I think at those times, the process for understanding what the members would need and value weren't very sophisticated. A lot of the governance process hadn't existed yet. In those early times, I think we were guessing about rules and functionality. Then, that turned into a voting process. During the TRAC certification times it was guesswork.” This interviewee was from a different repository than Repository Staff 18 above, whose organization did not have a mission statement.

Standard developers, auditors, and repository staff agreed that organizational instability was a potential source of risk for digital repositories. While standard developers and auditors agreed that a clear mission statement supported by well-documented policies would offset potential threats to repositories and digital content by ensuring that the repository maintained a focus on the goal of long-term preservation, repository staff were skeptical about the effectiveness of this type of documentation to offset these potential threats. Indeed, repository staff members who were interviewed said that they were able to provide the necessary documentation to achieve certification despite the fact that their repositories lacked the governance structures that they knew the standard was meant to enforce. In the case of repository documentation such as a mission statement, the difference between standard developers and auditors on one hand, and repository staff on the other, was in part a difference in understanding about the function of those policies. Repository staff members disagreed with the standard developers and auditors about whether the policies should reflect repository practices, or should describe aspirational best practices. Repository staff characterized such policies as ideals that their organizations thought they should meet, but also described their repositories as organizations that were shaped by power struggles and lacking in the social mechanisms needed to meet those ideals. This can be interpreted, in part, as a difference between trustworthiness and a desire to be trustworthy.

Communication, complexity, expertise, organizations, political culture, and vulnerability were all factors from my theoretical model that surfaced in relation to organizational governance as a potential source of risk for digital repositories. Interviewees described organizational governance structures as complex systems in which individuals and organizations with particular types of knowledge and expertise come together in order to carry out the work required for long-

term preservation of digital content. Although standard developers and auditors described documentation about repository policies as a way for repositories to mitigate threats relating to organizational instability, repository staff were skeptical about whether such documentation could stabilize relationships with more powerful institutions.

#### **4.3.4 Repository Processes**

Interviewees identified processes for digital object management as potential sources of risk for digital repositories and digital content. They discussed the ways that metadata creation, file format management, and processes, such as content ingest threatened the longevity of digital content as well as the ability of digital repositories to carry out their mission of long-term preservation. Auditors agreed with the view of risk presented by the standard developers, but repository staff argued that the actual work of managing digital content over time was not as straightforward as the TRAC standard implies. Repository staff described this section of the TRAC standard as one that generated the most disagreement with auditors during their audits, but staff members were able to sufficiently communicate their practices and policies, and the reasoning behind them, to obtain certification.

Standard developers discussed repository processes for digital object management, such as ingest, transformations, capture/creation and management of metadata, and content delivery as potential sources of risk for digital repositories and digital content. Through Section 4 of the TRAC standard, “Digital Object Management,” this view of repository processes as a potential source of risk is communicated to auditors and repository staff. Sub-sections covering ingest, preservation, management, and access of digital content make clear that potential threats exist throughout the entire lifecycle of a digital object, and suggest that repositories can demonstrate that they have sufficiently identified and addressed those threats through evidence contained in



documents including, but not limited to, policies, procedures, agreements, workflows, logs, and correspondence.

Standard developers described the goal of digital object management in a TDR as “selecting and preserving the information in a way that will be useful in a long-term preservation, as part of the long-term preservation goal” (StandardDeveloper\_01). This interviewee further explained that digital object management in the context of TRAC was about more than “just managing digital formats” (StandardDeveloper\_01). Rather, digital object management in the context of OAIS and TRAC was “concerned about preserving the information content, not just the format” (StandardDeveloper\_01).

Metadata creation, capture, and maintenance were discussed by Standard Developers 01, 04, 08, and 09. These standard developers explained that it was important for repositories to understand their Designated Communities in order to know what type of representation information they would need to capture in order to preserve digital content for future use, “if it's a long-term repository, preservation for the long-term, then the greatest risk is understanding what needs to be captured now so that the data can be understood in the future” (StandardDeveloper\_04). This interviewee went on to explain that lack of understanding about how important metadata are for long-term preservation was a threat to the long-term viability of digital content:

*“But all the other areas of which primarily I work in, for example, what metadata they have, whether it's representational information or context information, which is necessary for the use of data, oftentimes was ignored. The assumption was that if you handed somebody a spreadsheet they'd be able to figure it out. And you know that's probably true for active archives. But again, if you're talking about long-term archives, you know, 30 years from now if people get a spreadsheet they'd have no idea how to use it. And of course it gets worse. I mean, if it's one of some type of data structure that it would take a programmer to parse it and display it, if there's not information about what it means, it's basically useless. Sometimes this stuff is almost cryptography trying to figure out what it really*

*meant in the first place. If it's not documented then you can't figure it out. I think that's the pattern I found, or I personally observed, was the lack of real understanding of what it takes in terms of metadata to make something useful.”*  
(StandardDeveloper\_04)

For developers of the TRAC standard, having sufficient, appropriate metadata was crucial for long-term preservation of digital content, and this emphasis on representation information was reinforced through the standard itself, with several criteria in Section 4 asking for information about, and evidence of, metadata capture, creation, and maintenance in order to communicate that the repository has identified and addressed this potential threat.

Another common theme among standard developers was the challenge that file formats posed to long-term preservation of digital content, “the more formats that you are taking in and using for your AIPs, the more complex that gets, the combinatorics when you start talking about multiple file formats, multiple record types, compound records, software dependence of the records” (StandardDeveloper\_03). Standard Developers 03, 04, 06, and 08 all discussed potential threats to repositories and digital content relating to file formats, including potential problems, such as file format obsolescence, difficulties in sufficiently documenting unusual file formats, and the fact that repositories frequently lack the expertise, staffing, and funding to sustain the amount of work that would be necessary to support a large number of different file formats within one repository:

*“I think most archives have preferred formats and then they have other formats that don't get the support that they need. I think there's a lot of problems with long-term missions and funding for archives. That there's a lot of interest in data access, but less interest in the long-term preservation of data that, especially data that you don't know, whether there's some things that you'll need long-term and other things that only get used sporadically. In many archives there's such a low access to some pieces of their data, but they'll likely become very important in the future. And you find out they haven't been properly preserved in the meantime.”*  
(StandardDeveloper\_08).

Standard Developers 04, 05, 07, 09, and 10 identified repository actions around processes to ingest, migrate, and store digital content, as well as actions to verify the fixity or integrity of that content as potential sources of risk, “the fixity or the integrity of the data is critical” (StandardDeveloper\_04). Indeed, Standard Developer 05 explained that a number of factors during the ingest process that could negatively impact the repository and/or the longevity of the digital content:

*“You have to maintain, as much as possible, the control of what is going to be transformed. Some properties [have] to be transformed. And of course in this case you can accept the transformation. You must accept. Because the digital preservation is dynamic. Formats change, digital signatures cannot be verified. So you have to build a documentation system able to document which kind of transformations have been done, on which basis. Because many of [these] transformation[s] are not reversible. They are forever. You have change and you are going to lose the original things and what was.” (StandardDeveloper\_05)*

The standard developers presented a view of repository processes for digital object management as one that required substantial amounts of documentation in order to ensure that future custodians and users of digital content would be able to access and understand that content. While standard developers focused on the potential threat to digital content posed by repository staff failing to understand what information to capture, create, and maintain, I found that auditors were more concerned that even when repository staff knew what policies and practices they should have, repositories would lack the staffing, expertise, funding, or organizational will to carry out that work.

Interviewees described the work of digital object management as something that takes place across different functional areas of a repository, and explained that coordinating and managing this work was difficult, “In terms of the actual getting the work done from ingest to storage to metadata to access and all that, those functions can be spread all across the organization, whatever kind of organization they are. Being able to coordinate those functions

and have clear lines of authority about when a policy is put in place, who has to adhere to it, and where the responsibility lies, that can be very difficult to do.” (Auditor\_01)

Auditor 05 argued that repository processes for digital object management were a potential source of risk because of the likelihood that they would be abandoned or scaled back over time, “they start off with the goal of having defined processes, workflows, and all that sort of stuff, and over time a lot of that stuff gets either dropped or the period between things like migration activities or even just repository auditing activities expands as the organizations are pressed for resources and staff.”

Similarly, Auditor 06 described the processes for digital object management over time as having uncertain consequences over time, and therefore creating uncertainty about the longevity of digital content:

*“Digital preservation is a series of handoffs, probably every five to seven years. That you have to continually be touching, and curating, and evaluating content and digital collections or else they really will just die. And so the content is constantly being moved from server to server, and service provider to service provider, and operating among servers with various new operating systems, and all of that has effects that we can’t always predict. And so there’s just a whole lot of uncertainty around this whole notion of digital preservation that is new and different from [physical collections].” (Auditor\_06)*

In terms of errors that could occur in these processes, auditors argued that the stakes were high for repositories that focused on long-term preservation because of the likelihood that errors would go unnoticed for very long periods of time. For example, one interviewee described human error in repository processes for long-term preservation of digital objects as the greatest threat to digital repositories:

*“I think human failure, or failure in human-driven processes, which include a lot of technical processes. I mean, technical processes are only as good as the humans that develop them. The reason that I would distinguish between a preservation repository and a general access repository is if something goes wrong, if our Sys Admin makes a mistake or an error, introduces an error in an access repository, you’re likely to know about it right away. Whereas, if a human*

*introduces an error into a preservation repository, you may never know about it until it's too late. Until someone 50 years in the future goes to extract that AIP and gets something out of it and there's nothing in it.” (Auditor\_09)*

Auditors were concerned with issues relating to file formats, and argued that the understandability of different file formats over the long-term was a potential threat to digital content:

*“I think that one of the biggest challenges is to get back again to file formats, and which data structures have the potential for the longest understandability and usability, especially with science, and social science databases. Whether that be small surveys, or large, complex, relational databases. You know, how are we going to continue to be able to use those over time, and how will we be able to, if we will be able to, recover them in the future and make them understandable? And I think people are working on this, but I think it is probably the biggest challenge, in terms of digital object management” (A10).*

Overall, auditors understood the view of risk provided by standard developers through the TRAC standard, and agreed that repository processes for digital object management were a potential source of risk for digital repositories and the content that they were preserving. Yet, auditors were more focused on how lack of human resources and human error or loss of resources would impact a repository’s ability to carry out the processes necessary for long-term preservation, while standard developers were concerned about whether repositories would understand the needs of their Designated Communities well enough to capture appropriate representation information for preservation and reuse, and whether their workflows and procedures were comprehensive enough to capture all of the actions applied to their collections over time. Standard developers assumed that addressing this potential source of risk was a matter of having enough information and technical knowledge about digital object management, while auditors questioned whether that information was knowable and argued that it would not be possible over the short term to assess whether a repository’s digital object management processes were successful.

As with the standard developers and auditors, repository staff members also focused on metadata, file formats, and repository processes for digital object management as potential sources of risk for digital repositories. Repository staff identified metadata as an area that could pose a potential threat to both the repository and the digital content. Repository Staff 03 explained that poor metadata management practices could negatively impact the usefulness of a repository for its users, “The devil’s in the details. You can maintain preservation metadata and do it well. Or you could do it poorly. And so risks, I guess, implicit there are if it’s not normalized, if it’s not taking advantage of controlled vocabularies or authority, things like that, then the quality of the preservation metadata, if it’s poor, could present a risk to the usefulness of the repository” (RepositoryStaff\_03). In addition to maintaining metadata over time, Repository Staff 07 emphasized that metadata objects change over time and it is important for repositories to keep pace with the changes to digital objects and their metadata in order to preserve digital content, “in terms of the actual content itself what we're finding is that it all changes, and in particular the metadata about objects changes a lot more than the underlying objects themselves. Both in terms of being enriched and enhanced over time, but also in terms of just being corrected.” (RepositoryStaff\_07).

Repository staff agreed with the standard developers that the work of maintaining file formats over time could pose a potential risk to repositories because of the amount of time and resources required to do that work, “Those that do take heterogeneous content, yet have an expectation of format migration or access or even possibly normalization, but they'll take in heterogeneous content, anything where it's not as simple as possible, like I think that is tremendous complexity on all levels. I mean, sure, there's code to deal with it, and that can be hard, but harder than that is policies that you can actually apply and get buy-in and everything at

all of the places where your data might be coming from and at the right point”

(RepositoryStaff\_06).

However, repository staff disagreed that file format obsolescence or lack of expertise would be a problem for repositories and argued instead that as long as there was sufficient interest and knowledge in the repository or its Designated Community they would be able to make sense of the digital content, “So much data is not human readable. I mean you can maybe look into a file, but it generally requires software. And data in a lot of ways is almost software itself, right? I mean it is in a complex file format can be extremely difficult to parse. We can do it now, we have the tools and all that. And if you have enough ... if there's enough of something in the world, it's probably going to be all right.” (RepositoryStaff\_05). Repository Staff 04 pointed to current successes with outdated formats as an example, “You know, we've worried a lot in the preservation community about Word Perfect is gone. We can't read Word Perfect anymore or these weird file formats are gone and it's actually never been the case. We've never not been able to figure out what we've got, as long as we've still got it.” (RepositoryStaff\_04).

Repository Staff 08, 12, and 15 spoke at length about processes to ingest content as costly and time consuming, “Ingest of content is the most expensive piece, and it is where almost all the resources are spent. And unfortunately, true for [repository], true for everyone, the content that is most at risk is the most expensive to ingest.” (RepositoryStaff\_12). Similarly, Repository Staff 08 explained that it was costly to ingest digital content in a way that would support her repository’s mission of long-term preservation, “More often, however, the data would come to [repository] that had not been very rigorously produced or managed, and so it was expensive and time-consuming for us to process it in a way that allowed us to be confident of our preservation commitment.” (RepositoryStaff\_08).

Repository staff painted a picture of digital object management processes as ongoing, time-consuming activities that required regular actions but did not guarantee long-term success. Repository Staff 07 explained that digital content requires regular attention in order to ensure the integrity of each item, and also to ensure that it will be usable for the Designated Community, “There's so many items that can simply become obsolete as well as physically degrade and long-term digital preservation requires handling the data on a regular basis, so that you actually are continually testing your assumptions that it's not only still there but still usable and fit for a particular purpose.” On the other hand, Repository Staff 11 argued that the practical work of carrying out digital object management processes required making compromises in order to balance the other priorities of the repository against the amount of time and resources that would be needed to meet the ideal standards for digital object management processes, “One of the interesting things about being a preservation organization is that on the one hand you often have very high lofty ideals, but you have to balance that. There's a risk to meeting them. You have to balance that with the practical decisions” (RepositoryStaff\_11).

In contrast to the attitudes expressed by standard developers and auditors, that it was difficult for repository staff to meet the criteria set forth in the TRAC standard for digital object management, and that the discrepancy between the ideal and what repositories were likely to be able to accomplish presented a potential threat to repositories and content. Repository Staff 07 explained that this was an area of risk for digital repositories because best practices for managing digital objects for long-term preservation have yet to be established, “It quickly gets mind numbingly complex and we've talked about it a lot and have not come to any really good future-proof answers that we're comfortable with in terms of identifying objects uniquely, and perpetually, and persistently.” (RepositoryStaff\_07).



This disagreement between repository staff and standard developers and auditors about whether meeting the criteria described in the standard would ensure the longevity of digital content surfaced during the audit of Repository Staff 04's repository, "That was an awful lot of give and take. Then, there were a lot of revisions we had to do in our technical section because of that. I don't mean this as an insult, but they wanted clean, formulaic answers, and there just weren't any" (RepositoryStaff\_04). He emphasized that the auditors, following the TRAC standard, wanted his repository to provide clear responses to the criteria in Section 4, but that he found the actual work of managing digital objects to be complicated and messy. Indeed, several repository staff members identified this area as one where they disagreed with auditors, or where auditors required a substantial amount of additional information before they would agree to certify the repository.

Standard developers, auditors, and repository staff members all described processes for digital object management as a potential source of risk for repositories. While standard developers and auditors characterized digital object management as relatively straightforward and held that clear documentation of digital object management processes would mitigate risks in this area, repository staff argued that the actual work of managing digital content over time was not as straightforward as the TRAC standard implies. This was the section of the TRAC standard that repository staff reported as the most contentious during the audit process, because auditors wanted clear documentation communicating repository processes, and repository staff members viewed their processes for digital object management as complex and difficult to communicate via documentation.

Communication and complexity emerged as particularly strong factors in terms of repository processes for digital object management. All three groups, standard developers,

auditors, and repository staff, described digital object management processes as complex, but they disagreed about whether this complexity could be mitigated through clear documentation. While standard developers and auditors expected repository staff to communicate their processes in a clear, straightforward manner, repository staff members said that these processes were complex and messy, and did not lend themselves to simplification in the way that the audit process demanded.

#### **4.3.5 Technical Infrastructure**

Interviewees identified threats to the technical infrastructure of digital repositories as a potential source of risk. Standard developers and auditors described threats to technical infrastructure as identifiable and manageable and argued that repositories who engaged in the environmental monitoring required by the TRAC standard would be able to understand and respond to these threats. While some repository staff members agreed with this perspective, others questioned whether their repositories would be able to identify threats, and if they did identify them, thought that they might not have the resources to respond.

Among standard developers, threats to the technologies upon which repositories are built were described as a significant but manageable source of risk for digital repositories and their content. These interviewees identified aging hardware and software, costliness of maintenance, and the ongoing work required to sustain trustworthy infrastructure over time as potential sources of risk with straightforward solutions such as equipment replacement, software upgrades, content migration, and up-front investment in infrastructure.

Standard developers argued that the technical infrastructure of a repository was both complex and continually evolving as new digital preservation solutions come along, “The already complex world of hardware and software platforms. The concept of the virtual computer

has not proven to be very successful yet. We're stuck right now in, really, taking baby steps in terms of our hardware and our software approaches to digital preservation. We've got to get some kind of more universal, more virtual approach, to how we can preserve all formats of digital materials" (StandardDeveloper\_06). This complexity, they explained, requires continual monitoring in order to keep abreast of changes in the environment, "For example, one of the areas of the audit and certification standard is concerned with regular monitoring of changes in the environment, and that's complex because it can mean hardware obsolescence" (StandardDeveloper\_09).

The text of the TRAC standard echoes this belief in the importance of ongoing monitoring in order to maintain up-to-date hardware and software. Section 5, "Infrastructure and Security Risk Management" includes several criteria focused on monitoring different aspects of the repository infrastructure, such as 5.1.1.1 which states, "The repository shall employ technology watches or other technology monitoring notification systems" (Consultative Committee for Space Data Systems, 2012, p. 65). The supporting text for this item explains that, "This is necessary to track when hardware or software components will become obsolete and migration is needed to new infrastructure" (Consultative Committee for Space Data Systems, 2012, p. 65). Through this document, the standard developers frame threats to technical infrastructure as identifiable, often predictable, and as something that can be addressed before it becomes a problem.

While standard developers framed threats to technical infrastructure as manageable, they did point out that people were one of the biggest challenges in mitigating these threats, "The difficulty is always people. The hardware and software is always going to be much easier" (StandardDeveloper\_07). For example, while it might be relatively simple to set a timeframe for

hardware replacement, it may be difficult to secure the necessary funding to follow that replacement schedule, “You can't tell a resource allocator, as you suggested in here, which is a reasonable time frame, that you're going to basically wipe out everything and replace it all in three years. That what is brand new and spiffy and perfect now will all be gone in three years because it will be inadequate. Resource allocators don't like to hear that. They view this stuff as more permanent somehow. As more tangible assets, like a brick and mortar building, for example” (StandardDeveloper\_06).

In contrast, another standard developer argued that the cost of storage decreases exponentially over time, and that securing funding for long-term preservation was more about the ongoing work of digital object management rather than infrastructure:

*“So whereas initially [a] petabyte may be on one or two, maybe it's two tapes, in three years time it'll be on a small part of one tape. In another three years, it'll be on a tiny part of one tape, and in another three years it'll be next to nothing on a tape, and so the management of it will be negligible from then on because it's just this much of a tape and that's nothing in terms of the cost of the tape and the processing to check these things. So all of that is significant in terms of the costs, so then the costs come to actually making sure the data is usable.”*  
(StandardDeveloper\_07)

Standard developers framed threats to the technical infrastructure of a repository as ongoing and manageable. Standard developers established an understanding of long-term preservation in which digital content is expected to survive but the technologies used to store, preserve, and access it are not. Through the text of the TRAC standard, they communicated to both auditors and repository staff that a trustworthy digital repository should be able to demonstrate a firm understanding of the limitations of its infrastructure, and an ability to preserve digital content beyond the lifespan of any given part of that infrastructure.

Auditors were in agreement with standard developers about the importance of technical infrastructure and the notion that threats to technical infrastructure were significant but

manageable for digital repositories, “A technical infrastructure is not difficult. It may cost you a bunch of money, but it's a solvable problem and you kind of assume it's robust given that there are processes and checks and all sort of things in place to verify that it's robust” (Auditor\_09).

The expectation that money could solve problems relating to technical infrastructure was shared by several auditors, “Everything else from a, comes down to the challenge of technological change, but a lot of the technological change can be mitigated with sufficient resources” (Auditor\_05). Similarly, auditors argued that in addition to having sufficient resources, having appropriate staffing with the right kinds of expertise was also important for mitigating threats to the technical infrastructure of a repository, “The biggest thing I learned is that the human factors are more important than the technology factors. Because the technology factors, as long as you have good people and support for the technology, you can do that” (Auditor\_08).

Auditor 08 went on to explain that both the hardware and software of a repository require specialized knowledge and expertise, but that in general technologies for digital repositories are well known, “I think the technology is pretty well known at this point, and there's a variety of options that people can use in terms of the technology stack. There's a certain amount of specialized skills and knowledge that an organization needs to recruit for its core staff to be successful in the technology implementation around this, but you can usually find those people” (Auditor\_08). Implicit in this perspective is the assumption that with enough resources and the right kind of expertise, potential sources of risk to a repository’s technical infrastructure can be ameliorated.

In the context of a TRAC audit, one auditor explained that an important goal of the site visit is to inspect the physical infrastructure, including equipment, software, and facilities in

order to confirm that the documentation provided by repository staff accurately represents the repository, “You're there to gather evidence of facts, so yes, there is a data center and its doors are locked and under alarm. There is earthquake monitoring. So you know, one responsibility was to see things, okay? And I think that's really important. You see staff, you see equipment, you see servers, you're shown auditing software, and audit reports, and system logs, and all kinds of things. You see them live. So you're bringing evidence yourself, you're a witness”

(Auditor\_10). Auditor 08 described some of the questions that he has asked himself when assessing a repository's technical infrastructure during an audit: “Looking at the software itself, what's the nature of that? Is it proprietary commercial software? Is it community access, open-source software? What is the strength of the community around that software whether it's commercial or open-source, and to what extent does the organization engage with whatever that community is? Active engagement shows that the repository organization is close to the further development of that technology, which is something you want to see” (Auditor\_08).

One auditor explained that in his view, questions pertaining to the technical infrastructure of digital repositories will become easier to manage over time, “That the technical infrastructure is going to continue to evolve and the questions, I am imagining, will become more tractable over time” (Auditor\_04). In contrast, Auditor 05 said that the difficulty of predicting changes to technical infrastructure was a potential threat to digital repositories:

*“Probably from a complexity standpoint it's forecasting that technological change. There's a lot of expectation since there's fairly rapid change and you build your IDs around that, only to turn out that things don't change as quickly as you thought they were, which then ... I mean in theory that's a good thing, but it means that it's often hard to predict how much effort has to be done which then means that when you go back and you're requesting funding for the next cycle, the next year, the next whatever and you're asked, ‘Well, you didn't do that,’ and you say, ‘Well, because it wasn't necessary. There were no formats of significance that changed over that time period.’ Then when you come back and ask for that again and it's like, ‘Well, why do you keep including it if it's not happening?’*

*Conversely, sometimes things happen extraordinarily rapidly, especially in an emerging area, and no matter how much you budget for, or no matter how much you've planned for it, the speed of change exceeds what you've forecasted. It's really the complexity of identifying where those change inflection points are, where it can have an impact in terms of loss of data, in terms of inaccessibility of the existing data, that kind of stuff. I don't know if there's anyone who's actually any good at it.” (A05)*

Overall, auditors were in agreement with the view communicated by standard developers that although threats to the technical infrastructure of a repository were serious, they were also knowable and manageable. Both standard developers and auditors developed a view of potential sources of risk in this area as issues that repositories seeking TRAC certification should be able to identify and mitigate.

While repository staff members were in agreement with standard developers and auditors that threats to technical infrastructure were potential sources of risk for digital repositories, repository staff expressed mixed attitudes about the manageability of those threats. Some repository staff members agreed with the view of technical infrastructure as a potential source of risk that was manageable while others argued that technical issues could not be separated from other aspects of repository management, such as funding and staffing, and that problems in those areas had the potential to make threats to technical infrastructure intractable.

Repository staff who described threats to technical infrastructure as identifiable and manageable framed those issues as complicated but solvable, “Technical problems can all be solved” (RepositoryStaff\_07). For example, one interviewee explained that technical solutions, including responses to scaling issues, were things that a repository could handle, “I feel like there are complications on a technical level that you can just kind of deal with. Scaling issues are scaling issues, but ultimately you can deal with those with technical solutions” (RepositoryStaff\_018).

In contrast, several repository staff members described examples from their own experience where staffing issues compounded potential sources of risk relating to the technical infrastructure of their repositories. For example, one interviewee explained that staffing issues, including turnover, created instances where repository software was not understood by repository staff, “[We have] various generations of software and they've been developed by different people. We're not a huge organization, obviously, so it's not that big a deal, but we certainly have pieces of software that people are like, I have no idea what that is. Or, I know what that is, but I didn't write it. So I think that's really where most of our complexity lies” (RepositoryStaff\_04).

Another interviewee described the stakes for not understanding repository software as particularly high in instances where repository staff think that they understand their infrastructure and fail to catch problems until it is too late, “So that's a vulnerability. Especially software. You think it's doing one thing. Everybody thinks it's doing one thing, and then you find out if it's doing something else, and then maybe it's too late. I can't say I've seen that happen, but the potential is there” (RepositoryStaff\_05).

Repository Staff 03, an IT manager from a TRAC certified repository, described his approach to managing technical infrastructure as being driven by a desire to prevent the repository from being affected by a failure, “As far as the technical infrastructure too, I never wanted us to be impacted by failures. I never wanted to say, ‘We had some sort of system failure but, we think everything's okay.’ Or ‘Service was down for this time because of some unplanned thing that we didn't understand.’ I really tried to keep everything to a high bar in terms of those kinds of technical considerations. Redundancy for all of the – Also, I didn't want to respond to crisis. I didn't want my staff to have to respond to crises. You know?”



In addition to questioning whether breakdowns in technical infrastructure would be identified, repository staff also argued that repositories could not assume that they would always have the staffing to support their infrastructure and respond to potential threats, “We have three now. But we're still doing the work that we did when we were seven. So there's things that are not happening that I wish were happening. You know, even on a systems side, I haven't done these quick operating system upgrades and security patches as we did when we had the staff that was dedicated to that. So even though I'm the systems lead, I spend more time writing software than managing the systems, just because the software changes are more strategically important than some of the other things” (RepositoryStaff\_16).

While most of the interviewees focused on the functionality of the technical infrastructure of repositories as crucial for digital object management and long-term preservation of digital content, Repository Staff 04 argued that in his experience hardware and software were fragile but data were robust, “It definitely influences it on the technical side, because I've seen computers fail for 30 something years. They all failed and they all break. They're exceedingly fragile. Storage is exceedingly fragile, so that's always in the back of my mind for all those things. At the same time, I'm sounding like a broken record, but what has actually been very robust has been the data” (RepositoryStaff\_04). While most of the risks described in this section were framed by interviewees in terms of the longevity of digital content, interviewees tended to discuss technical infrastructure on its own terms.

Interviewees largely identified threats to the technical infrastructure of repositories as a potential source of risk that was straightforward and within the power of repository staff to address. While repository staff shared the understanding of this potential source of risk as communicated through the TRAC standard, they disagreed about whether responding to threats

in this area would be as clear-cut for their repositories as the standard developers and auditors assumed it would be.

The factors from my theoretical model for the social construction of risk that arose most frequently in this section were complexity and expertise. Standard developers, auditors, and repository staff members all described technical infrastructure as complex, they disagreed about whether repository staff would be able to identify and respond to threats to digital repositories. While some repository staff members agreed with standard developers and auditors in their characterization of technical infrastructure as manageable, other repository staff members argued that other areas of repository management such as funding and staffing would prevent their repository from maintaining the level of expertise needed to identify and mitigate threats to their technical infrastructure.

#### **4.4 TRAC Audit Process and Factors that Influence the Social Construction of Risk**

##### **4.4.1 Site Visit**

Standard developers, auditors, and repository staff described the site visit as an important element of the overall audit process. While standard developers and auditors characterized the site visit as an opportunity to provide information lacking in a repository's documentation, and to investigate whether the documentation accurately depicts a repository's day-to-day operations, repository staff described the site visit as an opportunity to confirm the accuracy of their documentation. All three groups emphasized the importance of face-to-face communication between auditors and repository staff members, and reported that the site visit allowed auditors to develop a better understanding of the repository as an organization.

Although the TRAC standard does not provide specific guidance about how the audit process should be conducted, several standard developers said that auditors should conduct a site

visit as part of their assessment. Standard Developers 01, 03, 08, and 10 all described a process whereby the audit team would review documentation provided by a repository, and then visit the repository in order to determine whether the documentation was an accurate depiction of the repository's day-to-day practices. For example, Standard Developer 03 explained that the face-to-face communication during the site visit was crucial for developing a shared understanding between auditors and repository staff about the checklist requirements and whether the repository staff had sufficiently addressed them, "But then often what you find when you get on-site is, you can develop sort of a picture of what a repository is doing from that documentation, but when you're actually there with the folks in front of you and you actually sit down with them and have them demonstrate some of the processes that they use you find new questions that you hadn't thought of that you need to ask them about."

Indeed, Standard Developer 08 also explained that the site visit provided an opportunity for auditors to gather information that would fill in gaps in a repository's documentation, "Then when you do the onsite audit portion of the audit, you sort of delve into those areas in a deeper way to-and try to talk with the staff and determine whether they just didn't have enough response they could provide at the time, or they didn't know the right things to put down. Sometimes you can help them get to a point where they know what's being asked and they can fill in the pieces."

Certification reports produced by CRL for each of the six TRAC certified repositories state that the audit process included a site visit by members of the audit team. Auditors explained that the activities that took place at each site visit varied, depending on the documentation submitted by the repository, "It really depends on what kinds of responses we get from the self-audit and the documentation what we decide to look at at the site visit, but we need to confirm all the things that we had questions about and talk to the right people to help clarify other questions.

Then we go back to the panel and talk to them about what we found, and then we come out with the report” (Auditor\_03).

Auditor 10 explained that the site visits generally included meetings and presentations as well as inspection of facilities, and that direct communication with repository staff was important for understanding the people and the organization, and putting the repository’s documentation into context, “you have staff meetings, you have presentations, tours of the ... you'd see all the hardware, and the servers. You'd check for security disaster plans. You meet with head of the IT, of the whole, usually with the person who's in charge of the data centers. Usually you meet with people from the organization” (Auditor\_10). She went on to explain that the act of observing repository practices in-person is an important part of the audit process, and that the site visit itself is part of the evidence that auditors should consider during an audit, along with the documentation, “You see staff, you see equipment, you see servers, you're shown auditing software, and audit reports, and system logs, and all kinds of things. You see them live. So you're bringing evidence yourself, you're a witness” (Auditor\_10).

Auditors 05 and 08 explained that while members of the audit advisory panel reviewed repository documentation, they did not participate in the site visit. Rather, a small team of auditors consulted with the advisory panel before and after the site visit, with the goal of developing a deeper understanding of repository policies and practices, and filling in gaps that the advisory panel had identified in repository documentation:

*“I think that in our case we were heavily dependent on the auditors to actually go to them and clarify some of the issues. Things like the site visit to actually verify that their documentation and their behavior were identical. I think for the most part the information that the auditors brought back was certainly at least verified most of what they were saying. Certainly in the documentation itself, in the repository's documentation, there was a number of things that were definitely, when we looked at it, more aspirational than actual. This is what they were*

*hoping to do, and they were planning to get there, type of thing. Certainly not all the stuff that we looked at were things that had materialized yet.” (Auditor\_05)*

Indeed, Auditor 08 explained that the advisory panel relied on auditors to use the site visit to further investigate problems that were identified through the documentation, “If we had really critical security issues that we brought up, those would translate into [auditor] and [auditor] following up when they did their site visit.”

Auditors described the site visits as a crucial element of the TRAC audit process, and stressed the importance of the information that the auditors gathered while on-site at a repository for their assessment of each repository, “you come to understand certain policies or decisions, when you're on site, that may be a little bit more difficult to grasp from paper” (Auditor\_10). In some cases, members of the audit advisory panel said that they would have liked to participate in the site visit, but acknowledged that it would be financially impractical, “There's probably things that I certainly would have liked to have been part of the site visits, for instance. Pragmatically, it would have been unrealistic from a cost perspective, but those are the kinds of things where you're aligning the documentation to actual behavior for instance” (Auditor\_05). It would also prevent the advisory panel from remaining anonymous during the audit process, “The panel never really talks to the repository” (Auditor\_03).

Repository staff agreed with the standard developers and auditors that the site visit was an important part of the TRAC audit process. While standard developers and auditors described the site visit as an opportunity to fill in gaps from the documentation, repository staff largely described the site visit as an exercise in confirming that their documentation was indeed an accurate representation of a repository's day-to-day operations, “my feeling about the on site was always that it was there just to make sure that everything from what they had been told up to that

point actually checked out. And to have some conversations in a slightly higher bandwidth environment” (RepositoryStaff\_18).

Repository Staff 07, 11, 13, 16, 18, and 19 all characterized the site visit as a positive experience. Repository Staff 11 and 13 both said that their site visits went smoothly because they had planned ahead in order to manage the process and keep both repository staff members and auditors focused and on-task throughout the visit:

*“We managed the on-site review pretty carefully. We divided it up into sections that corresponded to the sections of the readable documentation structure, not the 16363 one. We prepared presentations and demos for each of those sections. We had two, in some cases three, run-throughs with the whole team for each of those sections. We had set up visits to the secure machine rooms where the hardware was. Things like that. We actually tried to, we tried to control the on-site review process as much as possible, and I think that ended up working well, because by the time, we gave them time in each of the sections for asking questions, and then there was time at the end for them to go away and think about what they'd seen and then come back with more questions and we had pretty much answered their questions by the time they went away and started to think about other questions to ask.” (RepositoryStaff\_13)*

Similarly, Repository Staff 16 described the site visit as consisting of both discussions and demonstrations, “It was that there was a series of questions, but there was also a bit of a show-and-tell.” Repository staff generally characterized the site visits as successful, and said that they felt that by the end the auditors understood their repositories, “I think they had a pretty good understanding by the end of the meeting ... There seemed to be a pretty good understanding of how we operated, both at the technology level and governance level” (RepositoryStaff\_19).

In contrast, Repository Staff 21 said that she found the site visit to be superficial. She said that she had expected the auditors to be more skeptical of the repository, and to seek out evidence to support the repository’s documentation rather than trusting repository staff members at their word, “I just felt, they're not really diving in deep enough. It's just asking surface

questions based off of the checklist criteria and not really going in depth in terms of evaluating the content” (RepositoryStaff\_21).

Overall, standard developers, auditors, and repository staff were in agreement about the importance of the site visit and the value of direct, face-to-face communication between auditors and repository staff in establishing a shared understanding of repository policies and practices. While standard developers and auditors described the site visit as an opportunity to ask questions, gather additional information, and verify the accuracy of repository documentation, repository staff described the site visit as a chance to demonstrate the accuracy of their documentation and provide auditors with an understanding of their organization in order to provide context for their documentation. Auditors, a group whose professional backgrounds tended to emphasize administration and management, described a focus on getting answers to their questions, while repository staff, a group with diverse professional experience but whose roles tended to be focused on hands-on work rather than repository administration, expressed mixed opinions about the depth of understanding that auditors were able to reach in just a couple of days and in some cases argued that maintaining strict control over the site visit agenda helped them to shape the process.

#### **4.4.2 Maintaining Certification**

Standard developers, auditors, and repository staff members agreed that repositories should undergo periodic reviews to maintain TRAC certification. However, they disagreed about how frequently such reviews should happen and who should initiate them. The TRAC standard does not provide much guidance about how repositories should maintain certification, and standard developers did not discuss recertification much, but many auditors and repository staff members were under the impression that the TRAC standard called for recertification every three

years. This recertification schedule was not supported by the CRL certification reports, nor by the auditors who were directly involved in helping repositories to maintain their certification. Rather, auditors expected that repository staff members would contact them any time a new disclosure and/or inspection was needed, while repository staff members expected that their recertification reviews would be initiated by the CRL auditors.

The TRAC standard states that maintaining certification is an ongoing activity, “Attaining trustworthy status is not a one-time accomplishment, achieved and forgotten. To retain trustworthy status, a repository will need to undertake a regular cycle of audit and/or certification” (Consultative Committee for Space Data Systems, 2012, p. 19). While this does not specify a schedule, one standard developer explained that certification should entail annual audits as well as recertification every three years, “You have continuing audits. You have yearly maintenance audits and you have to get recertified ... every three years” (StandardDeveloper\_08).

CRL issued a certification report for each repository which specified that ongoing certification was contingent upon regular disclosures every two years as well as periodic inspection as determined by mutual agreement (Center for Research Libraries, 2010, 2011, 2012, 2012, 2013, 2014, 2015). Interviews with auditors revealed a lack of consensus among one another as well as with the text of the audit reports about how repositories should maintain their certification. Auditor 03 explained that, “As long as their conditions remain about the same as when they were certified, the TRAC certification holds. It doesn't have an expiration” (Auditor\_03). She added that repositories should notify CRL of any substantial changes, “They're supposed to tell us if there's any significant changes within the organization, and then we would decide whether or not we needed to review it” (Auditor\_03).



In contrast, Auditor 06 said that the repositories were told that they would be reviewed for recertification after three years, but noted that the three-year reviews have not happened, “The findings were issued generally with provisos that they would be revisited in three years. That hasn’t happened and I doubt it ever will happen.” Auditors 08 and 10 both agreed that repositories should undergo regular audits in order to maintain certification, but were unsure about what the schedule should be, “There’s some expectation that once you get audited, you renew that audit, that certification, with a subsequent audit. I think one question in the community is how often that should occur” (Auditor\_08).

While repository staff shared similarly mixed opinions about the audit schedule for recertification, most expected that CRL would initiate the process and several indicated that their repositories expected to be contacted for their recertification audit. For example, Repository Staff 07 said that his repository expected to be revisited by the CRL auditors every two years but at the time of the interview had yet to be contacted, “According to the certification report itself they require biennial disclosures so the first one wouldn't be due until June. We do internal reviews annually where we go through all the documentation process and update them. We did one internally last year but according to the conditions of certification they require disclosures every two years, so they haven't asked us yet. I'm assuming they will in another month or so but we haven't heard anything.” He said that his repository might follow up if they were not contacted by CRL, but that he was doubtful about the likelihood of being revisited, “I think if we didn't hear anything from them for an extended period of time then we might follow-up and ask. We also understand that this is not something that they have a tremendous amount of resources to devote to either. I’m not even sure if, it sounds like they might not want to be doing a lot of certifications anymore because they found them to be quite onerous.”

Similarly, Repository Staff 16 said that he thought his repository was due for a recertification review, “We have a renewal coming up.” Repository Staff 04 said that although his repository was told that they would have to go through recertification audits every three years, he thought that was an impractical timeframe in light of how long the first audit took, “When we were going through it, the time period that kept coming up was three years. You should get re-certified every three years. There were two issues with that. One is if you take a year and a half to certify us, taking every three years, we would never stop being audited.” He also noted that the three-year mark had come and gone with no word from CRL about recertification, “CRL has never come back to us about recertification.”

Several repository staff members said that their repositories were updating documentation in preparation for their recertification audit, including Repository Staff 16 who said that his repository was preparing for another audit despite being unsure what would be required of them for recertification, “The impression that I have of the update is that they may ask us questions, but they also may not ask us anything. So we’re updating our documentation just as we have been. So the idea is that the documentation isn’t supposed to be static. Documentation is supposed to be updated to reflect what is current at the time. And so a number of those processes have been updated since then. But whether they’ll be reading the documentation and asking us questions, I don’t know” (RepositoryStaff\_16).

In contrast, Repository Staff 11 and 17, who were from different repositories, were the only interviewees to report that their repositories were in regular contact with CRL in order to advise them of major changes. Repository Staff 11 said that her repository would reach out to CRL as needed to update documentation, “We actually stay in touch now. As things change or shift, we’ll reach out to CRL and let them know what we’ve got going on.” Repository Staff 17

was the only repository staff member who described a process of updating documentation and notifying CRL of those updates that matched the process described by Auditor 03, “We do an annual review that we submit to them. Sometimes they follow up, sometimes they don't. We did one last year and that just goes through all of the documentation, anything that's changed, we make note of it. So if our standards have been updated, or infrastructure had been changed, anything like that. Or staffing. So that's what I'm in charge of this year, so I'm doing TDR review for them, it's due at the end of the summer ... I'm not exactly sure if we send it or if we just do the review and then if they sometimes they ask us for it, and sometimes they don't. So it's more of an open ended process, but we just do it anyway.”

The open-ended statement in the TRAC standard about a regular cycle of audits to maintain certification left the actual schedule of review up to the auditing organization, and the certification reports indicate that the recertification schedule was left open at the end of the certification process as well. The lack of clarity about how repositories should maintain their certification and whether CRL would initiate a recertification process after a set period of time created a great deal of uncertainty among both auditors and repository staff about the certification process, and about the length of time that their certification would last, “If someone, anyone, came and said you should take that off your website, you should stop saying you're certified, I'd have a hard time arguing with them. If CRL came and said your period has ended, you're no longer certified, I probably would not disagree with them” (RepositoryStaff\_04).

As with the audit checklist and the anonymity of TRAC auditors, communication was a significant factor in this area. Audit reports indicate that an ongoing review schedule would be determined by mutual agreement between CRL and each repository (Center for Research Libraries, 2010, 2011, 2012, 2013, 2014, 2015), but most repository staff members said that the

email they received with the audit results was the last communication that they had with the audit team. With the exception of Auditor 03 and Repository Staff 11 and 17, neither auditors nor repository staff understood when or how recertification would happen. Even other staff members at the same repositories at Repository Staff 11 and 17 did not share their understanding of recertification.

Repository staff members described their organizations as continually changing to meet the needs of their stakeholders, and to keep up with new technologies. Most repository staff members who were interviewed for this study said that their repositories had implemented substantial changes since they achieved TRAC certification. TRAC is a certification system focused on demonstrating trustworthiness of digital repositories through a process of external review (Consultative Committee for Space Data Systems, 2012). As such, repositories claiming certification but operating under policies and practices that have not been reviewed – even when repository staff believe them to be improvements over the versions that were reviewed during their audit – calls into question whether external stakeholders can trust claims of TRAC certification to tell them anything about the current state of a repository.

#### **4.5 Conclusion**

My dissertation has demonstrated that individuals involved in the TRAC audit process largely understood risk in terms of specific threats to repositories and their digital content. Although they described risk as knowable and calculable and expressed the belief that people would behave in rational and predictable ways in response to risk information, my analysis revealed differences in how each of the three participant groups understood potential sources of risk. While standard developers, auditors, and repository staff tended to agree on the major categories of potential risk for digital repositories (i.e., finance, organizational governance, legal,

repository processes, and technical infrastructure), repository staff often disagreed with standard developers and auditors about whether the audit process could accurately assess their ability to mitigate those risks and ensure the long-term preservation of digital content. For example, repository staff, a group consisting of people with varied educational and professional experiences who tended to be less senior than the standard developers or auditors but were more likely to be in professional roles where they were directly carrying out the work of preserving digital content, expressed greater skepticism about the effectiveness of succession plans as mitigation tools for financial risk and their discussion about this topic tended to focus on the immediacy of the threat to their organization, their role, and the digital content that they were preserving.

Communication, expertise, uncertainty, and vulnerability were particularly strong factors that influenced how auditors and repository staff members understood risk in the context of TRAC audit processes. In particular, my findings suggest that these two groups were more likely to agree about the process and aims of the auditor site visit, in which auditors were able to interact and communicate directly with repository staff than the process for maintaining certification, which relied on limited written communication. As with the potential sources of risk identified above, repository staff members, who had higher levels of vulnerability or exposure to potential risks, were more likely to express skepticism about whether the evidential requirements outlined in the TRAC standard and enforced by auditors would ensure the long-term preservation of their digital content. They expressed doubt about whether the audit process could facilitate a thorough understanding of the complexities of their organizations, and concern about the expertise of the auditors.

Despite these differences of opinion, repository staff understood the requirements of the TRAC standard as communicated by standard developers and enforced by auditors, and produced the evidence that auditors wanted to see in order to achieve certification. This performance of trustworthiness enabled them to achieve their goal of becoming certified as a trustworthy digital repository, even when they themselves did not believe that their processes were trustworthy or likely to result in long-term preservation of digital content.

## **Chapter 5: Discussion & Conclusion**

### **5.1 Summary of Findings**

In the previous chapter I examined how standard developers, auditors, and repository staff involved in the TRAC audit process understood the concept of risk for digital preservation. I traced the social construction of risk through the eyes of three stakeholder groups involved in the TRAC audit process and found that even though the digital preservation community has relied on a classical definition of risk and assumed that people behave in a rational and predictable way in response to risk information, the results of this study show that repository staff members disagreed with standard developers and auditors about whether the risk identification and mitigation strategies prescribed in the TRAC standard would translate to actual trustworthiness with regard to long-term digital preservation.

This study was motivated by the following research questions:

1. How do standard developers, auditors, and repository managers conceptualize risk in the context of a TRAC audit?
2. What are the differences and similarities by which standard developers, auditors, and repository managers understand risk as it has been communicated by the TRAC standard?
  - a) In what ways do these differences and similarities become manifest in the TRAC audit process?

3. To what degree do the following eight factors which influence risk perception come into play in the audit process: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability?
  - a) In what ways and why do they emerge when staff and auditors consider risk factors articulated in the TRAC standard?
  - b) What additional factors, if any, emerge which also influence perceptions of risk in relation to the TRAC standard?

While standard developers and auditors believed that the requirements in a TRAC audit could assess a repository's ability to identify and mitigate potential risks in order to ensure the long-term preservation of digital content, repository staff often disagreed with standard developers and auditors about whether the audit process could accurately assess their ability to mitigate risk and ensure the long-term preservation of digital content. Repository staff members from TRAC certified repositories described meeting auditor expectations for certification across a variety of criteria, while at the same time disagreeing about whether the criteria themselves were an accurate measure of trustworthiness with regard to long-term digital preservation. The audit process itself provided the social context necessary for repository documentation to become evidence of trustworthiness (Amann & Knorr Cetina, 1988).

By examining two aspects of the TRAC audit process, auditor site visits and the process for maintaining certification, I found that communication, expertise, uncertainty, and vulnerability are factors that play a substantial role in the social construction of risk in the TRAC audit process. The other four factors from the model (i.e., complexity, organizations, political culture, and trust) were also present, but to a lesser degree. Findings from this study demonstrate that the factors in this model are not separate, but rather can be found in combinations that both amplify and minimize one another. For example, my findings show that poor communication about how repositories maintain their TRAC certification increased uncertainty about the audit



process itself for both auditors and repository staff members, and decreased trust between repository staff members and auditors.

## **5.2 Risk in Digital Preservation**

Risk is a foundational concept in digital preservation and the TRAC audit process (Consultative Committee for Space Data Systems, 2012; Conway, 1996). In Chapter 2, I characterized the classical definition of risk as one that included two elements that were common throughout the literature: (1) the probability, and (2) the magnitude of consequences of an event (e.g., Gardoni & Murphy, 2013; Hilgartner, 1992; Kaplan & Garrick, 1981; Leveson, Dulac, Marais, & Carroll, 2009; Rowe, 1977; Slovic, 1987). This understanding of risk relies upon the concept of a rational actor and assumes that different individuals will understand and respond to risk information in predictable ways. In this study, I found that the TRAC standard developers assumed that auditors and repository staff would interpret the TRAC standard in the same way, and that both groups would understand the risks facing a repository and agree on mitigation strategies and actions. Standard developers, auditors, and repository staff discussed the concept of risk in ways that demonstrated an understanding that reflected this classical definition. You will remember that in section 4.2, Repository Staff 18 and Auditor 10 both described risk as consisting of the probability and magnitude of consequences of an event. Both argued that while they did not believe that probability and magnitude of consequences were well known or understood in their field, that information could be known and people would behave predictably in response to it. Yet, the results of this research have shown that this understanding of risk was not reflected in their experiences. Instead, I found that individuals across those three groups did not share the same understanding of risk and did not agree about the risk mitigation strategies that were required for TRAC certification.

Interviewees described an audit process in which repository staff interpreted the requirements in the TRAC standard, and auditors evaluated the resulting documentation in order to enforce their own understanding of the requirements in the TRAC standard. My findings demonstrated that while interviewees from all three groups understood risk for digital repositories in relation to specific types of threats (i.e., finance, organizational governance, legal, repository processes, and technical infrastructure), standard developers and auditors, groups whose members described themselves in terms of their seniority, expertise, and leadership roles, believed that the TRAC audit process could accurately assess a repository's ability to mitigate those threats. Repository staff members, however, were skeptical. For example, section 4.3.1.1 demonstrated that standard developers and auditors believed that a succession plan was evidence that a repository had taken the appropriate steps to ensure that its digital content would survive if the repository were to fail. In contrast, repository staff members, a group whose members described themselves in terms of their functional roles within their repositories, did not believe that a succession plan was evidence that digital content would outlive a repository because they did not believe that the document would be followed or enforced.

My research demonstrates that identifying threats and describing mitigation techniques is a necessary step, but not sufficient in and of itself for digital preservation, because individuals have different understandings of risk and of whether and how risks can be mitigated. Although the classical definition of risk discussed above assumes that people will behave predictably in response to risk information, this study shows that individuals behave differently in response to risk information. In order to understand how digital repository stakeholders understand and respond to risk information, it is necessary to consider the social factors that influence and contribute to the construction of risk in TDR certification.

### **5.3 The Social Construction of Risk in TDR Audit & Certification**

The theoretical framework for this research argued that risk in digital preservation is socially constructed and that the following eight social factors influence the construction of risk in the TRAC audit process: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. My findings support the argument that digital preservation should treat risk as socially constructed phenomenon and consider how social factors contribute to an understanding of risk in the audit and certification of TDRs, and I found that communication, expertise, uncertainty, and vulnerability were particularly strong factors that influenced how auditors and repository staff members understood risk in the context of TRAC audit processes.

#### **5.3.1 Communication**

Theories of risk perception argue that risk is socially constructed and that communication can influence how individuals understand risk information, (e.g., Bostrom, 2014; Chung, 2011; Kasperson & Kasperson, 1996; Konheim, 1988; Lachlan, Burke, Spence, & Griffin, 2009; Renn, 1991; Renn, Burns, Kasperson, Kasperson, & Slovic, 1992). My findings show that communication was an important factor in risks related to finance, organizational governance, and repository processes. As you will recall, in section 4.3.1 repository staff members described communication with funders and parent organizations about the cost of long-term digital preservation as highly challenging. Similarly, in section 4.3.4 I found that repository staff viewed communication about repository processes for digital object management as complex and challenging, while standard developers and auditors thought that repository processes could be communicated in a clear, straightforward way.

Communication also came into play in the site visit (see section 4.4.1). Results from this study demonstrate that face-to-face communication was an important factor that enabled auditors and repository staff to establish a shared understanding of repository policies and practices during, for example, the site visit portion of a TRAC audit. In contrast, information that was not a part of the site visit, such as how to maintain certification, was communicated inconsistently. In section 4.4.2 I found that the TRAC standard, the audit reports, and individual interviewees all communicated different information about the process for repositories to maintain their certification and interviewees were largely unaware of these inconsistencies.

My findings emphasize the importance of communication throughout the TRAC audit process, and suggest that both auditors and repository staff members would benefit from a process that encouraged more direct communication between individuals in the two groups.

### **5.3.2 Expertise**

Research has demonstrated that individuals with different experiences and types of knowledge bring different types of expertise to bear on assessments of risk (Pidgeon, 1998), and that risk assessment processes should include individuals with different types of expertise and different perspectives (Slovic, 1987). This study shows that expertise was influential in how interviewees understood potential sources of risk relating to organizational governance and technical processes. For example, in section 4.3.3 expertise came up in relation to TRAC requirements that repositories have documentation about organizational structure and staffing to ensure that they have (1) appropriate staffing levels, and (2) have staff with sufficient expertise to carry out the work necessary for long-term digital preservation. While standard developers and auditors thought that the required documentation was evidence that a repository had the expertise necessary for the work of long-term preservation, repository staff said that there was a big

difference between understanding the types of expertise that was needed, and being able to sustain sufficient staffing levels to achieve that level of expertise.

Similarly, in section 4.3.5 repository staff disagreed with standard developers and auditors about whether threats to the technical infrastructure of a repository could be identified and mitigated. While standard developers and auditors described threats in this area as identifiable and manageable, repository staff members were skeptical about whether repositories could maintain the necessary staffing levels and expertise needed for that work.

As previously noted in section 4.4.1, expertise played an important role in the audit process during the site visit. For standard developers and auditors, the site visit was an opportunity to interact directly with repository staff who were experts in different aspects of repository management in order to gain a deeper understanding of repository policies and processes. For repository staff, the site visit was challenging because they had mixed opinions about whether the auditors who visited their repositories had sufficient expertise to understand their policies and processes in just a few days.

My findings indicate that standard developers, auditors, and repository staff view expertise as important for identifying and managing risks within repositories, and also for the TRAC certification process. While standard developers and auditors assumed that knowing what types of expertise were needed would ensure that repositories would be able to maintain that expertise on staff, repository staff thought that knowing what types of expertise were necessary for repository management was only the first step toward maintaining that expertise among a repository's staff.

### 5.3.3 Uncertainty

Research has shown that perceptions of risk can be influenced by the existence and recognition of uncertainty (e.g., Starr, 2003; van Est, Walhout, & Brom, 2012). This study found that uncertainty throughout the TRAC audit process influenced the ways that interviewees understood risk. For example, repository staff members were skeptical that the risk mitigation measures required by the TRAC standard, such as a succession plan, would ensure the long-term preservation of their digital content. This skepticism about the requirements for evidence in the TRAC standard extended across all five potential sources of risk (i.e., finance, organizational governance, legal, repository processes, and technical infrastructure), and demonstrated a different understanding of risk between standard developers and auditors on one hand, and repository staff members on the other.

For example, you will recall that in section 4.3.2, uncertainty came up when standard developers and auditors were more concerned with uncertainty about repository relationships among member and/or partner organizations as a potential source of risk for digital repositories, than repository staff members. Repository staff members discussed uncertainty about external organizations but were less likely to consider relationships with their partners as uncertain.

My findings demonstrate that uncertainty about the TRAC audit process was amplified by poor communication between auditors and repository staff members. As previously noted in section 4.4.2, information about how repositories should maintain their certification was communicated inconsistently, and as a result auditors and repository staff members expressed a great deal of uncertainty about what repositories should do to maintain their certification. The audit process did include some elements that were intended to improve communication and reduce uncertainty, such as the site visit (see section 4.4.1). Both auditors and repository staff

described the direct communication that the site visit facilitated as reducing uncertainty by providing an opportunity to fill in gaps in repository documentation and/or provide evidence for repository policies and processes.

These findings demonstrate that in areas where standard developers, auditors, and/or repository staff are aware of uncertainty, they took steps to reduce or eliminate it. However, standard developers and auditors appeared to be unaware that repository staff did not share their confidence in the risk mitigation strategies outlined in the TRAC standard. These differences suggest that standard developers, auditors, and repository staff members did not share the same understanding of the risks that mitigation strategies such as succession planning were intended to address.

#### **5.3.4 Vulnerability**

Research in the area of risk has shown that lived experience, including exposure or vulnerability to risk, can influence risk perception (e.g., Konheim, 1988; Olofsson et al., 2014). My findings showed that across all potential sources of risk, repository staff members were skeptical that the requirements of the TRAC standard would ensure the long-term preservation of their digital content. Standard developers and auditors described the goal of the TRAC audit process as a certification of a repository's ability to ensure the longevity of their digital content beyond the life of the repository itself.

This was quite salient in the area of financial risk (see section 4.3.1). Standard developers, auditors, and repository staff all described reliance on short term funding sources such as grants as increasing the vulnerability of a repository. Several repository staff members said that their repositories fell short of the requirements set forth by the TRAC standard to demonstrate mitigation of threats to financial sustainability, but all received TRAC certification.

Their views about whether their repositories had sufficiently addressed threats to financial sustainability were influenced by their own vulnerability to those risks. Repository staff members, who experienced greater vulnerability than either the standard developers or auditors to the potential sources of risk facing their repositories, were less likely to view those risks as manageable. Additionally, they did not believe that meeting the requirements described in the TRAC standard that the auditors were enforcing would make their repository trustworthy in terms of their ability to preserve digital content.

The TRAC audit process assumed that individuals would identify risks and agree about the appropriate mitigation techniques regardless of their closeness to the repository and/or digital content. Yet, findings from this study indicate that individuals outside of the repository, who may not have a strong understanding of a particular repository's policies and processes and whose livelihoods would not be threatened by repository failure, viewed risks as manageable while repository staff did not. As previously described in section 4.3.3, standard developers and auditors thought that a focus on long-term digital preservation, and a mission statement and policies that reinforced this focus, would together mitigate the potential risk of organizational instability. In contrast, repository staff members did not believe that policy documents would ensure that their repositories were able to maintain, for example, the necessary budget or staffing levels to carry out the policies in their documentation.

The differences between standard developers, auditors, and repository staff suggest that the TRAC audit process would benefit from increased communication between these groups in order to help auditors to understand the perspective of repository staff members who are vulnerable to the potential sources of risk that are detailed in the documentation prepared for their audit.



### 5.3.5 Complexity

Research shows that high levels of complexity in technical and social systems can make it difficult to identify probabilities, consequences, and hazards (e.g., Fischhoff, 1983; Perrow, 1999; Rijpma, 1997; van Est et al., 2012). My findings demonstrate that interviewees across all three groups (i.e., standard developers, auditors, and repository staff) viewed the organizational and legal structures governing digital repositories as highly complex, and that this complexity contributed to the construction of both organizational governance and legal issues as potential sources of risk. For example, you will recall that in section 4.3.3, Auditor 03 described governance and organizational stability issues as both complex and uncertain potential sources of risk for digital repositories.

In the context of a TRAC audit, interviewees found the organizational, legal, and technical aspects of repository management complex. While standard developers and auditors believed that the audit process was one that allowed auditors to understand and assess these complex policies and practices, repository staff disagreed. In their interviews, repository staff described the audit process as one in which they were able to provide auditors with the documentation that was needed to achieve certification, but described significant difficulties in communicating complex information to auditors. My findings also indicate that communication, uncertainty, and expertise interacted with complexity to both amplify and minimize potential sources of risk for interviewees. As previously discussed in section 4.4.1, repository staff members were skeptical about whether they could effectively communicate their repositories' complex policies and practices to auditors in a short amount of time. However, auditors and standard developers thought that the face-to-face communication that took place during the site

visit enabled them to gain a deeper understanding of complex repository policies and processes than they could achieve through email communication and documentation alone.

### **5.3.6 Organizations**

Each TRAC audit was ultimately a risk assessment of an organization. The team of auditors themselves represented an organization as well. Previous research has argued that risk analysis and risk management are both activities that take place within organizations, and these activities rely on social constructions of risk knowledge that are framed within the structure of the organization (Hutter & Power, 2005b). The findings from this study demonstrate that the standard developers, auditors, and repository staff involved in a TRAC audit represent different organizations that develop their own understandings of the concept of risk through the audit process. While standard developers and auditors tended to agree on their definitions of risk, repository staff members held opposing views about whether and how the risks that they identified during a TRAC audit could be mitigated. This difference in perspective largely fell across organizational lines – that is, the organization setting the standard and the organization enforcing the standard were in agreement about the effectiveness of the measures required, but members of the organizations being audited viewed the required risk mitigation measures as ineffective for long-term preservation.

Standard developer, auditor, and repository staff views of organizational instability as a potential source of risk for digital repositories were influenced by their views of the complexity of the repositories being audited and their documentation. You will recall that in section 4.3.3, standard developers focused on the ways that the requirements laid out in the TRAC standard would mitigate potential threats relating to organizational instability, auditors and repository staff

were skeptical about whether policies and documentation could accurately capture actual repository practices.

The difference in perspective between individuals being audited and those conducting the audit that this study has identified echoes findings from previous research about the ways that notions of risk are shaped within organizations, and the difficulties or disagreements that arise when those views are challenged by external actors such as auditors (e.g., Vaughan, 1996). In the case of a TRAC audit, the highly formalized communication process both facilitated and hindered communication between the two organizations, and the fact that repositories could achieve certification as trustworthy without believing themselves to be trustworthy highlights the fact that the TRAC process required repository staff to understand how standard developers and auditors view risk but did not require auditors to understand how repository staff viewed risk within their organizations.

### **5.3.7 Political Culture**

Research about risk and political culture has shown that individuals perceive risks within their own cultural and political context, and it is within this same context that decisions about how to respond to those risks are formulated and implemented (e.g., Dake, 1991; Jasanoff, 1986, 1998; Parthasarathy, 2007). Additionally, attitudes about risk reflect feelings of power, or lack of power, in relation to potential sources of risk (Jasanoff, 1998). In this study, I found that political culture, both in terms of cultural and political context, and feelings of power and/or powerlessness in relation to risk, influenced how standard developers, auditors, and repository staff understood risk in the context of a TRAC audit. As previously discussed in section 4.3.2, interviewees across all three groups (i.e., standard developers, auditors, and repository staff identified national context as a potential source of risk for the repository described in the

vignette. Specifically, they discussed potential problems for the repository that could arise as a result of having data storage locations in two different countries.

Standard developers and auditors tended to say that well documented agreements and/or contracts could mitigate threats in areas where repositories were subject to decisions of more powerful organizations, such as institutional support from funding organizations and/or parent institutions. However, you will recall that in section 4.3.1 and 4.3.2 repository staff members did not agree that contracts and agreements would be enforceable, and did not think that this type of documentation was evidence of their ability to preserve digital content.

### **5.3.8 Trust**

TRAC certification is a process by which digital repositories demonstrate their trustworthiness to preserve digital content long-term. Research about trust and risk has shown that information about risks cannot be separated from their contexts. Rather, trust (or mistrust) is an important factor in the relationships between individuals and institutions, and can be a guiding factor in how risk is defined (Nelkin, 1989; Wynne, 1992). In addition to being a foundational concept for the certification itself, trust between auditors and repository staff was a factor in the audit process.

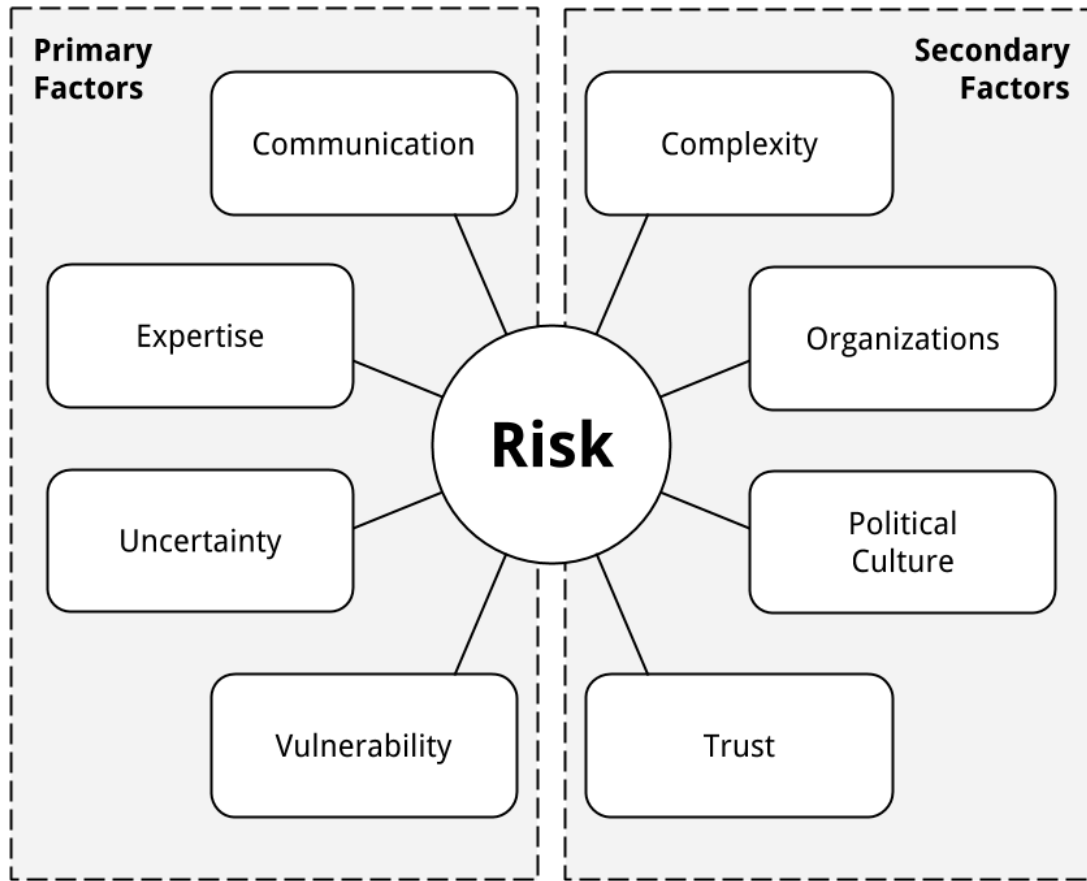
Through the course of a TRAC audit, repository staff members created documentation in response to the checklist in the TRAC standard, and auditors assessed and evaluated that documentation. The successful completion of each TRAC audit that interviewees discussed depended on trust between those two groups that repository staff were creating accurate and truthful documentation, and that the auditors were sufficiently expert in their areas of specialty to evaluate each repository's documentation. You will recall that in section 4.4.1, the site visit was described by both groups as an opportunity for repository staff to meet and speak directly with

auditors, and for auditors to confirm the accuracy of repository documentation. Trust but verify was a sentiment that ran through responses from auditors and repository staff, and direct communication was the means by which they verified trustworthiness.

As a foundational concept for TRAC certification, trust was notably problematic in light of my findings which show that repository staff used certification to perform the trustworthiness that they understood standard developers and auditors wanted to see, without believing that they were actually trustworthy with regard to long-term digital preservation. The succession plan is an example of this. In section 4.3.1.1, I found that repository staff provided evidence of their succession plans in order to demonstrate their repositories' ability to ensure the longevity of digital content beyond the lifespan of the repository, even though they did not believe that a succession plan provided evidence that the digital content would outlive the repository. While interviews with the standard developers, auditors, and repository staff showed that they trusted one another to be honest and accurate through the TRAC audit process, it is troubling that they disagreed about the assessment of trustworthiness at the heart of TRAC certification.

### **5.3.9 Revised Theoretical Model for the Social Construction of Risk in Digital Preservation**

In light of the findings described above, I have revised the theoretical model for the social construction of risk in digital preservation to reflect the fact that communication, expertise, uncertainty, and vulnerability emerged as particularly strong factors that influenced how auditors and repository staff members constructed their understanding of risk in the context of TRAC audit processes. In contrast, complexity, organizations, political culture, and trust also influenced the social construction of risk in the context of a TRAC audit, but to a lesser degree.



**Figure 4: Revised Theoretical Model for the Social Construction of Risk in Digital Preservation**

#### **5.4 Theoretical Contributions And Implications For Research, Policy, & Practice**

This research has brought empirical methods to an emerging discipline where scholarship has consisted mainly of case studies produced by practitioners. TDR certification is a relatively new phenomenon, and its reach is currently expanding beyond North America in part due to the recent adoption of ISO 16919 and the accreditation of the first organization to conduct ISO 16363 audits (i.e., Primary Trustworthy Digital Repository Authorisation Body Ltd.). As such, this research has created a set of baseline data about the first wave of TRAC certifications that will lay a foundation for future research.

#### **5.4.1 The Social Construction of Risk in TRAC Audit and Certification**

In this dissertation, I proposed a theoretical model for the social construction of risk in TDR certification that consists of eight factors: communication, complexity, expertise, organizations, political culture, trust, uncertainty, and vulnerability. The findings from this study have shown that each of the eight factors from the model contributed to the construction of risk in the TRAC audit and certification process, and that communication, expertise, uncertainty, and vulnerability were the most prominent factors.

Previous research has examined digital preservation as a technical, economic, and organizational phenomenon. The demonstrated applicability of this theoretical model to the TRAC certification process shows that risk cannot be examined solely as a discoverable, calculable phenomenon that relies on rational actors who behave predictably in response to risk information. Rather, my findings indicate that future research about digital preservation should consider it to be a social phenomenon. Risk is a foundational concept in digital preservation (Conway, 1996), yet it has been largely overlooked by research up to this point. This study has demonstrated that different actors in the TRAC certification process do not share the same understanding of the concept of risk. Research about digital preservation has largely treated risk as knowable and calculable and assumed that different people will behave predictably in response to risk information. These findings show that research about risk in digital preservation should consider it as a social construct and seek to understand whether and how different individuals and/or groups understand risk before proposing risk management solutions for digital repositories.

While this study did show that the factors from the model influenced how participants understood the concept of risk in the context of a TRAC audit, there was some difficulty in

differentiating social from individual factors. Vulnerability and expertise were factors in which social construction was the clearest. For example, you will recall from section 5.3.4 that repository staff members experienced greater vulnerability than either the standard developers or auditors to the potential sources of risk facing their repositories and were less likely to view those risks as manageable. When considering potential sources of risk for a repository in the context of a TRAC audit, there is more at stake for staff members than for the standard developers or auditors, for whom the audit process is more of a mental exercise without immediate personal threats. Interviewees from the group most likely to be directly effected by risks to repositories, the repository staff members, expressed skepticism about whether the risk identification and mitigation strategies described in the TRAC standard would translate to actual trustworthiness for their repositories. Interviewees from groups that were more removed from the repositories themselves, the standard developers and auditors, were more likely to view the TRAC criteria as sufficient measures of repository trustworthiness. Vulnerability is a factor that influenced how the three different groups (i.e., standard developers, auditors, and repository staff members), and individual interviewees, constructed their understanding of risk within the TRAC audit process. Future research should further disambiguate individual and social factors that influence how people and groups construct their understanding of risk in digital preservation.

#### **5.4.2 Repository Management & Trustworthy Digital Repository Certification**

My findings indicate that digital repositories can meet the requirements from the TRAC standard without the repository staff believing that those requirements will, in fact, ensure the longevity of their digital content. Standard developers and auditors were in agreement about the types of evidence that would demonstrate trustworthiness for digital repositories, but repository



staff members disagreed about whether documentation such as a succession plan was in fact evidence of repository trustworthiness.

Standard developers, a group consisting largely of individuals with graduate degrees in highly technical fields such as physics and engineering, established guidelines for repository certification that assumed identifying risks and describing policies and processes to address them could demonstrate a repository's ability to preserve digital content for the long-term. This approach to risk typified the shared epistemic culture among standard developers that emphasized discoverable, calculable phenomena rather than socially constructed phenomena and assumed that different people would behave rationally and predictably when presented with risk information. Similarly, the auditors enforced this understanding of how to determine repository trustworthiness. Their acceptance of the requirements set forth in the TRAC standard, and the underlying assumptions about risk identification and policy documents as sufficient evidence of repository trustworthiness, reflects the culture and expectations of this group of academic library administrators.

In contrast, repository staff members did not believe that documentation about repository policies and processes was evidence of trustworthiness with regard to long-term preservation of digital content. Rather, members of this group questioned (1) whether documentation would translate into action, and (2) if it did, whether those actions would produce consistent results. This group, which consisted of individuals with a variety of educational backgrounds and professional experience who were more likely than either the standard developers or auditors to be responsible for enacting the policies and processes described in TRAC documentation, did not believe that documentation was evidence of a repository's ability to preserve digital content long-term.

For example, you will recall in section 4.3.1.1 that standard developers and auditors believed that a succession plan was evidence that a repository's digital content would survive beyond the life of the repository itself, while repository staff members did not. The groups consisting of individuals with high levels of seniority and leadership experience in their professional roles expected that policies and agreements developed by repository leadership would be followed, while the more junior group of repository staff members were likely to argue that repository policies and agreements would not be followed or enforced. This example reinforces the differences between standard developers and auditors, and repository staff members, and highlights the ways that their different levels of education, expertise, and experience shaped their views of risk in the context of a TRAC audit.

In the future, repository staff should consider what measures and the corollary evidence they think would increase their perception of the trustworthiness of their repository with regard to long-term digital preservation and whether/how those measures complement or conflict with the accepted best practices for digital preservation and repository management. Rather than proceeding with certification under an evidential regime in a standard that they disagree with, the results of this research suggest that repository staff should take a more active part in the development of the standards themselves and that standard developers and auditors would benefit from including the perspectives of this group, which have so far been missing from the conversation.

In this research, I found that communication problems between auditors and repository staff contributed to the understandings of risk that each group constructed through the course of a TRAC audit. Because external parties such as standard developers and auditors have been shown to have different views from people within repositories regarding the effectiveness of risk

mitigation techniques, digital repository managers would do well to consider that the views that their staff members have of their repositories likely do not match the way that their repositories are viewed by standard developers, auditors, and other external stakeholders. Rather than maintaining these differing perspectives by maintaining one view of the repository internally and presenting another to external groups such as auditors, repository managers should consider increasing their involvement in the establishment of digital preservation policy and standards for TDRs.

Including the perspectives of internal and external stakeholders in the development of repository certification standards and digital preservation policy, including junior members of repository staff who can bring fresh perspectives and may have knowledge about new and emerging technologies, will strengthen both digital repositories as well as TDR certification processes.

### **5.4.3 Digital Preservation Policy**

Transparency is a key tenet of TRAC certification (Reilly, Jr. & Waltz, 2013), yet the findings from this study indicate that the audit process lacked transparency, “The TRAC audit process is, at least was, not transparent. It was intentionally not transparent” (RepositoryStaff\_04). For example, in section 4.4.2 standard developers, auditors, and repository staff all reported different processes for repositories to maintain their certification. The lack of consensus, and lack of clear communication about expectations among the three groups created uncertainty among auditors and repository staff about how repositories can remain trustworthy. Going forward, repository certification under the ISO 16363 standard should have clear and consistent guidelines for maintaining certification.

Auditors have presented repositories with reports of their findings, which are also made publicly available. Despite the length of the audit process and the detail of the TRAC checklist, these reports are brief statements about the overall status of the repository and the conditions at the time of the audit (Center for Research Libraries, 2010, 2011, 2012, 2013, 2014, 2015). A more transparent audit process would produce detailed results showing how the repository had met, or failed to meet, each of the checklist items. In contrast, the audit process examined in this study was one in which repositories would receive a score in each of the three areas of the checklist (i.e., organizational infrastructure, digital object management, infrastructure and security risk management), but the scoring system did not have a minimum score necessary to pass. Indeed, the scoring system for TRAC certification is such that while a repository can receive a low score, it is not possible to fail.

Increased transparency in the TRAC audit process would help bridge differing perceptions of risk among standard developers, auditors, and repository staff members. In addition to helping repository staff members understand the audit process itself, including how they should go about maintaining their certification over time, increasing transparency in the audit process would enable them to work with the auditors in order to find ways to meet both the letter and spirit of the TRAC standard. Transparency about the development of the standard itself could also provide opportunities for auditors and repository staff members to provide feedback about their experiences in order to improve the standard in the future.

## **5.5 Future Directions**

This research was motivated by (1) the lack of existing empirical research about digital preservation, and about TDR audit and certification processes in particular, and (2) the lack of research about the social construction of risk in TDR certification. This dissertation has

accomplished those goals and has produced a thorough investigation of the social construction of risk in the TRAC audit and certification process. However, more work is needed to understand how risk is constructed in other repository certification contexts. My future work will further refine the theoretical model that I have developed for the social construction of risk in digital preservation by applying it in different social, cultural, and organizational contexts where digital information is curated. This initial development and testing of my theoretical model has examined the social construction of risk in the context of large, well-resourced digital repositories in North America. Next steps for the development of this model include applying it to repositories that are at different stages of development and maturity, to repositories of varying size and resources, and in different cultural contexts. TRAC certified repositories represent an elite group, I would also like to apply this framework to a lighter certification such as the recently established CoreTrustSeal certification.

This study employed a qualitative mixed-methods research design that included in-depth semi-structured interviews and document analysis. These methods were appropriate here because of the exploratory nature of this research and the small size of the population. A larger study with more participants may offset some of the social desirability effects that were present in this research. A study with larger groups would also address the challenges encountered in this study of examining how groups construct an understanding of risk in contrast to individual perceptions of risk.

Alternately, an in-depth case study of a specific repository audit would address this study's limitations of limited recall and rationalization. Examining an audit as it takes place with ethnographic methods would provide a rich picture of how participants construct their understanding of risk, and how the factors from the model influence that process.

This research focused on three particular stakeholder groups: standard developers, auditors, and repository staff. Future research about repository certification should consider the perspectives of other repository stakeholders as well. For example, members of a repository's Board of Directors, or those in leadership roles within a repository's parent and/or member organizations may have different perspectives on the need for, and benefits of repository certification. Other stakeholder groups that may provide different perspectives on TDR certification include repository users, data depositors, clients/customers, and decision makers from funding organizations.

This study showed that the eight factors in the model for the social construction of risk influenced how interviewees understood risk in the context of a TRAC audit. However, there was some difficulty in differentiating between factors that influenced individual perceptions of risk and factors that influenced the social construction of risk. A study that focuses on the broader social context of repository certification, including other stakeholder groups as well as other data sources, could provide further insight into the question of what factors influence the social construction of risk in digital preservation, and which influence individual perceptions of risk.

Finally, the basis of this research can be traced back to the Garrett and Waters report, written in 1996, which called for the establishment of a standard for trustworthy digital repositories (Garrett & Waters, 1996). While this dissertation has focused on how stakeholders in the TRAC audit process construct their understanding of risk, a foundational concept for TDR certification, it has also shown that these groups have differing views about the effectiveness of the certification. This suggests that the time may be right for a study focusing on the value and

effectiveness of TDR certification, which considers whether and how the digital preservation community has taken up the recommendations from the Garrett and Waters report.

This dissertation sought to examine the social construction of risk in TDR certification. It was motivated by the need to examine the concept of risk, which is foundational for digital preservation but has not been examined thoroughly as a social phenomenon despite extensive research in other fields that demonstrate the importance of considering risk as a social construct. It has been successful in that it yielded data showing how standard developers, auditors, and repository staff involved in the TRAC audit process understood the concept of risk for digital preservation. This research traced the social construction of risk through the eyes of three stakeholder groups involved in the TRAC audit process and found that even though the digital preservation community has relied on a classical definition of risk and assumed that people behave in a rational and predictable way in response to the same information, the results of this study show that repository staff members disagreed with standard developers and auditors about whether the risk identification and mitigation strategies prescribed in the TRAC standard would translate to actual trustworthiness with regard to long-term digital preservation. Repository staff met the requirements for certification as trustworthy, but did not believe that these requirements were an accurate assessment of their ability to preserve digital content long-term. In doing so they were performing rather than demonstrating trustworthiness.

## **Appendices**



## Appendix A: Interview Protocol for Standard Developers

### Introduction Questions

1. Please tell me a bit about yourself and your background. How did you come to your current role? (OR, how came to be involved in PTAB)
  - a. NOTE: e.g., education, previous organizations, previous roles at current organization
2. How would you describe your current role at \_\_\_\_\_ (name of organization/repository)?
  - a. (NOTE: administration, IT, digital preservation, other)
  - b. What is your current job title? What was your title at the time of the TRAC audit?
  - c. How long have you worked at (name of organization/repository)? How long have you been in your current role?
3. Have you ever been an auditor for any repository assessment? (e.g., Data Seal)
4. Have you worked for a repository that went through a TRAC audit? (*goal is to find out if they've been on the other side of an audit.*)
  - a. An audit for any other certification?

### Vignette Questions

1. What risks, if any, do you see in the first category listed in the repository description (organizational infrastructure)?
2. In the second (digital object management)?
3. The third (infrastructure and security risk management)?
4. Of those risks, which is the most significant?
  - a. Why is this vulnerability significant?
  - b. What steps might Repository X take to address or mitigate this vulnerability?
  - c. What challenges do you think Repository X will encounter when they address these risks?

5. Is there anything about this repository that you think would be problematic if they decided to pursue TRAC certification?
  - a. What problems might arise?
  - b. How could the repository address these problems?

### **Interview Part 2: PTAB**

5. Please tell me a bit about yourself and your background. How did you come to be (a member of the PTAB board OR a participant in PTAB training)?
  - a. **For PTAB board**
    - i. What does your role as a PTAB board member entail? (training auditors, conducting audits, etc.)
    - ii. Have you been involved in training auditors?
      1. If yes, what is your role in the training process? (e.g., do you specialize in one area of repository assessment? one section of the TRAC checklist?)
    - iii. Were you involved in the 6 test audits that were conducted with 16363?
    - iv. Have you been an auditor for repository certification via the ANAB certification process?
  - b. **For PTAB trainees:** What did the training session entail?
    - i. How will you apply your training as an auditor?
      1. (e.g., conduct audits, support my own repository, consult with other repositories, etc.)

*For all PTAB: (use 16363 test audits if appropriate)*

1. What is the greatest risk or threat that digital repositories face?
  - a. Why is this risk significant?
2. What do you see as the area of greatest complexity for digital repositories? Why?
3. What are the most significant sources of uncertainty for digital repositories? Why?
4. In what ways does your experience [as an auditor] (or experience as \_\_\_role) influence your understanding of the risks that repositories face?
5. As an auditor, how do you identify risks/threats/vulnerabilities when you're conducting a repository audit?
  - a. How do you communicate information about them to the rest of the audit team?  
To the repository?
  - b. How is information about risks from other members of the audit team communicated to you?
6. How certain are you that you can rely on others in the audit team to tell you about risks that they have identified?

- a. Are they reluctant to talk about risk?

**Wrap-up**

- 6. Is there anything I haven't asked about that you would like to discuss?
- 7. Is there anyone else from your PTAB training session(s) who I should speak with?

## **Appendix B: Interview Protocol for Auditors**

### **Introduction Questions**

7. Please tell me a bit about yourself and your background. How did you come to your current role? OR How did you come to be involved in CRL's TRAC audits?
  - a. NOTE: e.g., education, previous organizations, and previous roles at current organization
8. How would you describe your role at (name of organization/repository)?
  - a. (NOTE: administration, IT, digital preservation, other)
  - b. What is your current job title? What was your title at the time of the TRAC audit?
  - c. How long have you worked at (name of organization/repository)? How long have you been in your current role?
9. Did you go through training in order to become a TRAC auditor?
  - a. Please describe any training activities, etc. that you participated in.
  - b. Have you been through PTAB training?
10. Have you ever been an auditor for any (other) repository assessment? (e.g., Data Seal)
11. Have you worked for a repository that went through a TRAC audit? (*goal is to find out if they've been on the other side of an audit.*)
  - b. An audit for any other certification?

### **Vignette Questions**

6. What risks, if any, do you see in the first category listed in the repository description (organizational infrastructure)?
7. In the second (digital object management)?
8. The third (infrastructure and security risk management)?
9. Of those risks, which is the most significant?
  - a. Why is this vulnerability significant?

- b. What steps might Repository X take to address or mitigate this vulnerability?
  - c. What challenges do you think Repository X will encounter when they address these risks?
- 10. Is there anything about this repository that you think would be problematic if they decided to pursue TRAC certification?
  - c. What problems might arise?
  - d. How could the repository address these problems?

## **Part 2: Auditor Experience**

*[To follow directly after the Vignette questions.]*

- 12. What is the greatest risk or threat that digital repositories face?
  - c. Why is this risk significant?
- 13. What do you see as the area of greatest complexity for digital repositories? Why?
- 14. What are the most significant sources of uncertainty for digital repositories? Why?
- 15. In what ways does your experience as an auditor (or experience as \_\_\_role) influence your understanding of the risks that repositories face?
- 16. As an auditor, how do you identify risks/threats/vulnerabilities when you're conducting a repository audit?
  - a. How do you communicate information about them to the rest of the audit team at CRL? To the repository?
  - b. How is information about risks from other members of the audit team communicated to you?
- 17. How certain are you that you can rely on others in the audit team to tell you about risks that they have identified?
  - a. Are they reluctant to talk about risk?

## **Part 3: Recent Audit Experience**

*The next set of questions asks about the most recent TRAC audit that the auditor conducted.*

- 18. During the TRAC audit process did you identify any previously unknown risks or vulnerabilities for your repository?
  - a. How were they identified?
  - b. How did you communicate them to the other auditors? To the repository?
- 19. Was there any disagreement among auditors about how to assess to the repository? Please describe.
  - a. How were they resolved?

20. Was there any disagreement between repository staff and the CRL auditors about the materials that the repository provided to you? Any relating to documents that you assessed/evaluated specifically? Please describe.
  - a. How were they resolved?
21. How certain are you that the information provided by the repository for the audit was complete and accurate?
  - a. Why/Please explain?
22. How did you communicate the audit results to the repository staff? (who, how, when, etc.)
23. Are there any red flags that you look for when auditing a repository?
  - a. Are there markers that you look for in order to identify areas of risk or vulnerability for repositories?
  - b. Are there certain things that you look for in the documentation that repositories provide? In the on-site visit?

**Wrap-up questions**

24. Is there anything I haven't asked about that you would like to discuss?
25. Who else in your organization should I speak with about the TRAC certification process?
  - a. NOTE: also include people who may have left the organization since then, where are they now?

## **Appendix C: Interview Protocol for Repository Staff**

### **Introduction Questions**

11. Please tell me a bit about yourself and your background. How did you come to your current role?
  - a. NOTE: e.g., education, previous organizations, and previous roles at current organization
12. How would you describe your role at (name of organization/repository)? (NOTE: if not in same role, ask about role at time of audit)
  - a. (NOTE: administration, IT, digital preservation, other)
  - b. What is your current job title? What was your title at the time of the TRAC audit?
  - c. How long have you worked at (name of organization/repository)? How long have you been in your current role?
13. Have you ever been an auditor for any repository assessment? (e.g., Data Seal)

### **Vignette Questions**

14. What risks, if any, do you see in the first category listed in the repository description (organizational infrastructure)?
15. In the second (digital object management)?
16. The third (infrastructure and security risk management)?
  - e. Of those risks, which is the most significant?
    - i. Why is this vulnerability significant?
    - ii. What steps might Repository X take to address or mitigate this vulnerability?
    - iii. What challenges do you think Repository X will encounter when they address these risks?

17. Is there anything about this repository that you think would be problematic if they decided to pursue TRAC certification?

f. What problems might arise?

g. How could the repository address these problems?

**Interview Part 2: repository staff, questions about risk**

18. What is the greatest risk or threat that your repository faces?

a. Why is this risk significant?

19. What do you see as the area of greatest complexity for your repository? Why?

20. What are the most significant sources of uncertainty for your repository? Why?

21. In what ways does your experience as \_\_\_\_ influence your understanding of the risks that your repository faces?

22. How do you identify risks/threats/vulnerabilities in your area/department?

a. How do you communicate information about them to the rest of the repository?

b. How is information about risks from other areas/departments communicated to you?

23. How certain are you that you can rely on others in your organization to tell you about risks that they have identified?

a. Are they reluctant to talk about risk?

**Interview Part 3: repository staff, audit process**

*The next set of questions asks about the TRAC audit process, including document preparation, the on-site visit, and any communication with auditors before, during, and after.*

24. During the TRAC audit process did you identify any previously unknown risks or vulnerabilities for your repository?

a. How were they identified?

b. How did you communicate them to the rest of the repository? To the auditors?

25. Was there any disagreement among repository staff about how to respond to the TRAC criteria checklist? Please describe.

a. How were they resolved?

26. Was there any disagreement between repository staff and the CRL auditors about the materials that your repository provided to them? Any relating to documents that you prepared specifically? Please describe.

a. How were they resolved?

27. How were audit results communicated to you? (who, how, when, etc.)

28. How certain were you that you could rely on the auditors to communicate important information back to your team at the repository during the audit process?



- a. How certain were you that you could rely on your team at the repository to communicate important information to the auditors?

**Wrap-up questions**

1. Is there anything I haven't asked about that you would like to discuss?
2. **Who else in your organization should I speak with about the TRAC certification process?**
  - i. **NOTE: also include people who may have left the organization since then, where are they now?**
  - ii. **ALSO: if interviewee interacted with auditors or any other CRL personnel, ask for their names**

## **Appendix D: Interview Protocol: Repository Staff, Audit Manager**

### **Introduction Questions**

29. Please tell me a bit about yourself and your background. How did you come to your current role?
- a. NOTE: e.g., education, previous organizations, and previous roles at current organization
30. How would you describe your role at (name of organization/repository)? (NOTE: if not in same role, ask about role at time of audit)
- a. (NOTE: administration, IT, digital preservation, other)
  - b. What is your current job title? What was your title at the time of the TRAC audit?
  - c. How long have you worked at (name of organization/repository)? How long have you been in your current role?
31. Have you ever been an auditor for any repository assessment? (e.g., Data Seal)

### **Vignette Questions**

32. What risks, if any, do you see in the first category listed in the repository description (organizational infrastructure)?
33. In the second (digital object management)?
34. The third (infrastructure and security risk management)?
- h. Of those risks, which is the most significant?
    - i. Why is this vulnerability significant?
    - ii. What steps might Repository X take to address or mitigate this vulnerability?
    - iii. What challenges do you think Repository X will encounter when they address these risks? Is there anything about this repository that you think would be problematic if they decided to pursue TRAC certification?

- i. What problems might arise?
- j. How could the repository address these problems?

**Interview Part 2: repository staff, questions about risk**

- 35. What is the greatest risk or threat that your repository faces?
  - a. Why is this risk significant?
- 36. What do you see as the area of greatest complexity for your repository? Why?
- 37. What are the most significant sources of uncertainty for your repository? Why?
- 38. In what ways does your experience as \_\_\_\_ influence your understanding of the risks that your repository faces?
- 39. How do you identify risks/threats/vulnerabilities in your area/department?
  - a. How do you communicate information about them to the rest of the repository?
  - b. How is information about risks from other areas/departments communicated to you?
- 40. How certain are you that you can rely on others in your organization to tell you about risks that they have identified?
  - a. Are they reluctant to talk about risk?

**Interview Part 3: repository staff, audit process**

*The next set of questions asks about the TRAC audit process, including document preparation, the on-site visit, and any communication with auditors before, during, and after.*

- 41. During the TRAC audit process did you identify any previously unknown risks or vulnerabilities for your repository?
  - a. How were they identified?
  - b. How did you communicate them to the rest of the repository? To the auditors?
- 42. Was there any disagreement among repository staff about how to respond to the TRAC criteria checklist? Please describe.
  - a. How were they resolved?
- 43. Was there any disagreement between repository staff and the CRL auditors about the materials that your repository provided to them? Any relating to documents that you prepared specifically? Please describe.
  - a. How were they resolved?
- 44. How were audit results communicated to you? (who, how, when, etc.)
- 45. How certain were you that you could rely on the auditors to communicate important information back to your team at the repository during the audit process?
  - a. How certain were you that you could rely on your team at the repository to communicate important information to the auditors?

#### **Interview Part 4: repository manager questions**

3. When did (name of repository) decide to pursue TRAC certification?
  - i. How was that decision made? Who initiated? Who were the decision-makers?
  - ii. How was this decision communicated to repository staff?
  - iii. Timeline:
    - i. How long did the entire process take from the decision to pursue certification until CRL posted the audit results?
    - ii. NOTE: follow-up for breakdown if necessary: decision to pursue until first contact with auditors; working with auditors and preparing materials for submission; auditors reviewing documentation & preparing decision
4. Please describe your role in the certification process.
  - iv. Which parts of the process were you involved in?
  - v. Which documents/evidence were you involved in preparing?
  - vi. Please describe any interaction that you had with the auditors.
    - i. Who did you communicate with? How? (i.e., phone, email, etc.)
  - vii. How did you coordinate activities among repository staff?
    - i. Did you encounter any difficulties while coordinating work among repository staff for the audit? Please describe.
5. Did your repository encounter any challenges or difficulties in preparing materials for the audit? Please describe.
  - viii. NOTE: follow-up for detail if necessary:
    - i. Internal disagreements
    - ii. Disagreements with auditors
    - iii. Problems interpreting and/or responding to the checklist
    - iv. Other challenges
  - ix. What were the primary reasons for these challenges? (e.g., communication problems, different perspectives about risk, lack of trust, uncertainty, etc.)
    - i. How did you overcome them?
6. Which areas of the TRAC checklist did repository staff spend the most time responding to for the audit?
  - i. Who was responsible for these areas?
  - ii. Why do you think that these areas took longer than others to address?
7. Which areas of the TRAC checklist do you think are the most important in terms of (your organization's) ability to preserve and provide access to digital information?
  - i. Which areas/items, if not addressed, would pose the greatest threat?

### **Wrap-up questions**

8. Is there anything I haven't asked about that you would like to discuss?
9. **Who else in your organization should I speak with about the TRAC certification process?**
  - iii. **NOTE: also include people who may have left the organization since then, where are they now?**
  - iv. **ALSO: if interviewee interacted with auditors or any other CRL personnel, ask for their names**

## Appendix E: Vignette

**Please read the repository overview below. During the interview, you will be asked to identify and discuss potential areas of risk for Repository X.**

### **Repository X: Overview**

#### Organizational Infrastructure

Founded in 2005, Repository X is a nonprofit organization whose mission focuses on both preserving and providing access to social science research data. Repository X was originally established as a partnership between eight large research universities located across the United States and Canada. Beginning in 2008 the repository expanded with a tiered dues-based membership system, and currently has 45 members. Repository X is managed by an Executive Director as well as a volunteer Board of Directors consisting of representatives from each of the original eight partner organizations as well as four external members who serve 2-year terms. Funding for Repository X comes from a combination of membership dues (~65%) and grant funding (~35%) as well as support from the original eight member organizations.

#### Digital Object Management

Current collections include 100 terabytes of data from both the original eight member organizations as well as the 45 member organizations. Repository X accepts data in any format but provides varying levels of support for different formats. Preferred file formats (e.g., .tiff, .wav, .csv, .txt) receive higher levels of preservation support, such as file migration, than those in other formats (e.g., .png, .mp3, .xls, .rtf), which receive bit-level preservation. 60% of current holdings are preferred formats. For those preferred formats, the repository maintains preservation metadata in addition to the metadata received from the data depositors.

### Infrastructure and Security Risk Management

One of the founding member universities hosts the repository infrastructure, and repository staff who manage day-to-day activities are based at this same university in the Pacific Northwest. Two active mirror backup sites are located in geographically diverse locations, one on the East Coast of the United States and another in Quebec, Canada. Repository X has a disaster response and recovery plan, which is updated every three years. The repository's preservation planning includes a succession plan.

## Appendix F: Interview Data Analysis Code Set

Parent Code	Child Code	Definition
accreditation	[none]	code discussion about accreditation for auditors to administer ISO 16363 (16919)
Audits	[none]	n/a
Audits	formal audit	code when discussion is about a specific audit that was conducted with the goal of certification
Audits	self assessment	code discussion about repository conducting self assessment
Audits	test audit	discussion about test audits conducted by CRL and PTAB
designated community	[none]	code discussion about the concept of designated community - they must use the specific term
documentation	[none]	discussion about documents or documentation in the TRAC audit process, documentaiton of repository policies, processes, procedures, etc.
good quotes	[none]	use to capture any short quotes or phrases that are particularly good
governance	[none]	use to capture discussion of repository governance
Interaction	[none]	use code (and all child codes) for any discussion of interaction/communication - face-to-face, telephone, email, etc.
Interaction	among audit team	
Interaction	among repository staff	
Interaction	between repository staff and auditors	
Most Significant Risk	[none]	code Q&A for question: what is the most signification risk that digital repositories face? OR what is the most significant risk that your repository faces/faced?
Organizations	[none]	n/a
Organizations	ANSI anab	
Organizations	CCSDS	code when participant discusses the Consultative Committee for Space Data Systems



Organizations	CLIR	
Organizations	CRL	code when participant discusses CRL
Organizations	IMLS	code when participant discusses IMLS
Organizations	JISC	
Organizations	LOCKSS	code when participant discusses LOCKSS
Organizations	NASA	code when participant discusses NASA
Organizations	PTAB	code when participant discusses PTAB
Participant Background	[none]	n/a
Participant Background	Education	discussion of participant's education background
Participant Background	Other Assessment	code any mention of participant experience with other assessments, either as auditor or working at a repository that pursued the assessment
Participant Background	Work Experience	discussion of participant's work experience
PTAB training	[none]	use when participants talk about the PTAB training sessions
Publishing	[none]	any time participant talks about publications
Repositories	[none]	for parent code & child codes, apply to all mentions of repository (not just first mention)
Repositories	Canadiana	any discussion of Canadiana, especially specific mentions of the repository
Repositories	Chronopolis	any discussion of Chronopolis, especially specific mentions of the repository
Repositories	CLOCKSS	any discussion of CLOCKSS, especially specific mentions of the repository
Repositories	HathiTrust	any discussion of HathTrust, especially specific mentions of the repository
Repositories	Portico	any discussion of Portico, especially specific mentions of the repository (sometimes but rarely referred to as Jstor or ITHAKA)
Repositories	ScholarsPortal	any discussion of ScholarsPortal, especially specific mentions of the repository; may also be called OCUL
Risk Definition	[none]	code if participant explicitly defines risk or gives clear statement of what they consider risk to be with regard to digital preservation or digital repositories
Risk Factors	[none]	social factors that influence participant risk perception
Risk Factors	communication	Perceptions of risk vary depending on the way in which information about those risks is communicated, including the source, method, channel, and means of communication. These elements can either amplify or attenuate perceptions of risk for different individuals and groups (e.g., Bostrom, 2014; Kasperson & Kasperson, 1996).
Risk Factors	complexity	High levels of complexity can make identification difficult with regard to hazards, probabilities, and consequences. Complexity in systems can also lead to unexpected interactions between component parts, often leading to increased levels of risk (e.g., Perrow, 1999; Wilkinson, 2001).

Risk Factors	expertise	Both expertise and lack of expertise can influence perceptions of risk. Experts may have particular knowledge that allows them to understand risk in a particular area, but they have been found to have a narrow focus based on their specialized knowledge, which can influence their perception of risk. Individuals who lack expertise in a particular area may not have the same nuanced understanding of particular areas that experts do, but they have been found to have a greater sense of the broad social context within which they are operating (e.g., E. Vaughan & Seifert, 1992; Wynne, 1992).
Risk Factors	organizations	Organizations both produce and manage risk, and perceptions of risk vary for people depending on their position within an organization. Risk assessment and management activities take place within the context of organizations, and are therefore influenced by the organizations themselves as well as the roles of the individuals within the organizations who participate in those activities (e.g., Hutter, 2005; D. Vaughan, 1996).
Risk Factors	political culture	National context influences how risks are defined. Perceptions of risk are shaped not only by the political culture within which individuals exist, but also by their place or role within that culture. These factors can elevate or reduce perceptions of risk depending on the position of an individual within the culture. Decisions about how to manage and respond to risks are shaped by political culture as well (e.g., Dake, 1991; Jasanoff, 1986).
Risk Factors	trust	Organizations and processes that involve cooperation by people and groups with different types of knowledge and expertise require trust among those actors. Perceptions of risk can vary depending on the amount of trust that these individuals and groups have for one another (e.g., Nelkin, 1989; Wildavsky & Dake, 1990).
Risk Factors	uncertainty	In many situations it can be difficult to determine and understand risk and its components (hazard, probability, consequences). People and groups operating under conditions of uncertainty may perceive risks differently depending on their level of uncertainty (e.g., Starr, 2003; van Est et al., 2012).
Risk Factors	vulnerability	Risk exposure, or vulnerability, influences perceptions of risk. People and groups who are able to limit their risk exposure may have different perceptions about risk than those who lack the ability to manage their exposure to risks. Greater vulnerability has been shown to increase perceptions of risk, while privilege and the ability to limit or select risk exposure has been shown to decrease perceptions of the severity of risks (e.g., Murphy, 2006; Olofsson et al., 2014).
Risk Sources	[none]	code specific sources of risk discussed by participant
Risk Sources	financial	discussion of financial risks or finances as a source of risk
Risk Sources	legal	discussion of legal risks
Risk Sources	organizational	discussion of organizational risks or organizations as a source of risk
Risk Sources	people	discussion of people as site or source of risk
Risk Sources	processes	discussion of processes, especially preservation processes, as site or source of risk

Risk Sources	technology	discussion of technology as risky or as a source of risk
Specific Certifications	[none]	n/a
Specific Certifications	DRAMBORA	Any mention of DRAMBORA
Specific Certifications	DSA	Any mention of Data Seal of Approval
Specific Certifications	ISO 14721	Any mention of OAIS or the ISO standard
Specific Certifications	ISO 16363	Any mention of the ISO standard
Specific Certifications	ISO 16919	Any mention of the ISO standard
Specific Certifications	ISO 27000	Any mention of the ISO standard
Specific Certifications	nestor	
Specific Certifications	TRAC	Any mention of TRAC
TDR Certification	[none]	
TDR Certification	Attitudes	participant's attitude toward certification
TDR Certification	Audit Outcomes	code discussion of audit outcomes
TDR Certification	Audit Process	code discussion of the audit process itself
TDR Certification	Benefits	discussion of benefits of certification
TDR Certification	Challenges	discussion of challenges for certification
TDR Certification	Site Visit	discussion about the site visit of the audit
Vignette	[none]	use parent code to block code entire set of questions about vignette
Vignette	2 Digital Object Management	code participant discussion of risks identified in the digital object management section of the vignette (2nd section)
Vignette	3 Infrastructure and Risk Management	code participant discussion of risks identified in the infrastructure section of the vignette (3rd section)
Vignette	1 Organizational Infrastructure	code participant discussion of risks identified in the organizational infrastructure section of the vignette (1st section)
Vignette	Significant Risk	code participant discussion of most significant risk for vignette

## References

- 4C Project. (2016, April 25). About 4C [Nonprofit]. Retrieved April 25, 2016, from <http://4cproject.eu/about-us>
- About Chronopolis. (2016). [Education]. Retrieved April 7, 2016, from <https://libraries.ucsd.edu/chronopolis/about/index.html>
- About HathiTrust. (2016). [Nonprofit]. Retrieved April 7, 2016, from <https://www.hathitrust.org/about>
- Anderson, J. C., & Gerbing, D. W. (1991). Predicting the Performance of Measures in a Confirmatory Factor Analysis with a Pretest Assessment of Their Substantive Validities. *Journal of Applied Psychology*, 76(5), 732–740. <https://doi.org/10.1037/0021-9010.76.5.732>
- Arvai, J. L. (2007). Rethinking of Risk Communication: Lessons from the Decision Sciences. *Tree Genetics & Genomes*, 3(2), 173–185. <http://doi.org/10.1007/s11295-006-0068-7>
- Babbie, E. R. (2010). *The Practice of Social Research*. Belmont, CA: Wadsworth Cengage.
- Baker, M., Shah, M., Rosenthal, D. S. H., Roussopoulos, M., Maniatis, P., Giuli, T., & Bungale, P. (2006). A Fresh Look at the Reliability of Long-term Digital Storage. In *Proceedings of the 1st ACM SIGOPS/EuroSys European Conference on Computer Systems 2006* (pp. 221–234). New York, NY, USA: ACM. <http://doi.org/10.1145/1217935.1217957>
- Ball, A. (2010). *Review of the State of the Art of the Digital Curation of Research Data*. Bath, UK: University of Bath. Retrieved from <http://opus.bath.ac.uk/18774/2/erim1rep091103ab11.pdf>
- Barateiro, J., & Borbinha, J. (2012). Managing Risk Data: From Spreadsheets to Information Systems. In *Electrotechnical Conference (MELECON), 2012*. (pp. 673–676). <http://doi.org/10.1109/MELCON.2012.6196521>
- Barateiro, J., Antunes, G., & Borbinha, J. (2011). Long-Term Security of Digital Information: Assessment Through Risk Management and Enterprise Architecture. In *2011 IEEE EUROCON - International Conference on Computer as a Tool (EUROCON)* (pp. 1–4). <http://doi.org/10.1109/EUROCON.2011.5929270>

- Barateiro, J., Antunes, G., & Borbinha, J. (2012). Manage Risks through the Enterprise Architecture. In 2012 45th Hawaii International Conference on System Science (HICSS) (pp. 3297–3306). <http://doi.org/10.1109/HICSS.2012.419>
- Barateiro, J., Antunes, G., Freitas, F., & Borbinha, J. (2010). Designing Digital Preservation Solutions: A Risk Management-Based Approach. *International Journal of Digital Curation*, 5(1), 4–17. <http://doi.org/10.2218/ijdc.v5i1.140>
- Beck, U. (1992). *Risk Society: Towards a New Modernity*. London; Newbury Park, CA: Sage Publications.
- Beck, U. (1999). *World Risk Society*. Malden, MA: Polity Press.
- Becker, C., & Rauber, A. (2011). Decision Criteria in Digital Preservation: What to Measure and How. *Journal of the American Society for Information Science and Technology*, 62(6), 1009–1028. <http://doi.org/10.1002/asi.21527>
- Berman, F. (2008). Got Data?: A Guide to Data Preservation in the Information Age. *Communications of the ACM*, 51(12), 50–56. <http://doi.org/10.1145/1409360.1409376>
- Berman, F., Kozbial, A., McDonald, R. H., & Schottlaender, B. E. C. (2008). The Need to Formalize Trust Relationships in Digital Repositories. *Educause Review*, 43(3), 11–12.
- Berman, F., Lavoie, B., Ayris, P., Choudhury, G. S., Cohen, E., Courant, P. N., ... Van Camp, A. (2010). *Sustainable Economics for a Digital Planet: Ensuring Long-Term Access to Digital Information*; Blue Ribbon Task Force on Sustainable Digital Preservation and Access Final Report.
- Bernard, H. R. (2012). *Social Research Methods: Qualitative and Quantitative Approaches*. Sage Publications, Incorporated.
- Beyond Google Books: Getting Locally-Digitized Material into HathiTrust. (2015, June 18). [Nonprofit]. Retrieved April 7, 2016, from <https://www.hathitrust.org/blogs/perspectives-from-hathitrust/beyond-google-books-getting-locally-digitized-material-hathitrust>
- Bostrom, A. (2014). Progress in Risk Communication Since the 1989 NRC Report: Response to “Four Questions for Risk Communication” by Roger Kasperson. *Journal of Risk Research*, 1–6. <http://doi.org/10.1080/13669877.2014.923032>
- Bradley, K. (2007). Defining Digital Sustainability. *Library Trends*, 56(1), 148–163. <http://doi.org/10.1353/lib.2007.0044>
- Bryce, H. J. (2007). The Public’s Trust in Nonprofit Organizations: The Role of Relationship Marketing and Management. *California Management Review*, 49(4), 112.
- Canadiana.org. (2015). About Canadiana.org | Canadiana [Nonprofit]. Retrieved March 30, 2016, from <http://www.canadiana.ca/en/about>

CBC News. (2012, May 2). Federal Libraries, Archives Shutting Down [News]. Retrieved June 19, 2016, from <http://www.cbc.ca/news/canada/ottawa/federal-libraries-archives-shutting-down-1.1139085>

Center for Research Libraries. (2010). CRL Certification Report on Portico Audit Findings.

Center for Research Libraries. (2011). CRL Certification Report on the HathiTrust Digital Repository.

Center for Research Libraries. (2012). CRL Certification Report on Chronopolis Audit Findings.

Center for Research Libraries. (2013). CRL Certification Report on Scholars Portal Audit Findings.

Center for Research Libraries. (2014). CRL Certification Report on CLOCKSS Audit Findings.

Center for Research Libraries. (2015). CRL Certification Report on the Canadiana.org Digital Repository.

Charmaz, K. (2006). *Constructing Grounded Theory*. London ; Thousand Oaks, CA: Sage Publications.

Chronopolis Homepage. (2016). [Education]. Retrieved April 7, 2016, from <https://libraries.ucsd.edu/chronopolis/>

Chronopolis Pricing. (2016). [Education]. Retrieved April 7, 2016, from

Chung, I. J. (2011). Social Amplification of Risk in the Internet Environment. *Risk Analysis*, 31(12), 1883–1896. <http://doi.org/10.1111/j.1539-6924.2011.01623.x>

Clifton, G. (2005). Risk and the Preservation Management of Digital Collections. *International Preservation News*, (36), 21–23.

CLOCKSS. (2014, July 28). CLOCKSS Archive Certified as Trusted Digital Repository; Garners top score in Technologies... [Nonprofit]. Retrieved March 30, 2016, from <https://www.clockss.org/clockss/News>

CLOCKSS. (2015). FAQ - CLOCKSS [Nonprofit]. Retrieved March 30, 2016, from <https://www.clockss.org/clockss/FAQ>

Colati, J. B., & Colati, G. C. (2009). A Place for Safekeeping: Ensuring Responsibility, Trust, and Goodness in the Alliance Digital Repository. *Library & Archival Security*, 22(2), 141–155. <http://doi.org/10.1080/01960070902904118>

Consultative Committee for Space Data Systems, Data Seal of Approval Board, & DIN Working Group “Trusted Archives - Certification.” (n.d.). *European Framework for Audit and Certification of Digital Repositories*. Retrieved August 11, 2014, from <http://www.trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html>

Consultative Committee for Space Data Systems. (2011). Space Data and Information Transfer Systems — Requirement for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories (Standard No. ISO/DIS 16919). Washington, D.C.: Consultative Committee for Space Data Systems.

Consultative Committee for Space Data Systems. (2012). Space Data and Information Transfer Systems — Audit and Certification of Trustworthy Digital Repositories (Standard No. ISO 16363:2012 (CCSDS 652-R-1)). Washington, D.C.: Consultative Committee for Space Data Systems.

Consultative Committee for Space Data Systems. (2012a). Reference Model for an Open Archival Information System (OAIS) (Magenta Book No. CCSDS 650.0-M-2). Washington, D.C.: Consultative Committee for Space Data Systems.

Contribute to CLOCKSS. (2015). [Nonprofit]. Retrieved April 7, 2016, from [https://www.clockss.org/clockss/Contribute\\_to\\_CLOCKSS](https://www.clockss.org/clockss/Contribute_to_CLOCKSS)

Conway, P. (1996). *Preservation in the Digital World*. Washington, D.C.: Commission on Preservation and Access.

Corritore, C. L., Kracher, B., & Wiedenbeck, S. (2003). On-Line Trust: Concepts, Evolving Themes, a Model. *International Journal of Human-Computer Studies*, 58(6), 737–758. [http://doi.org/10.1016/S1071-5819\(03\)00041-7](http://doi.org/10.1016/S1071-5819(03)00041-7)

Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Los Angeles, CA: Sage.

Creswell, J. W. (2013). *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Los Angeles, CA: SAGE Publications.

Dake, K. (1991). Orienting Dispositions in the Perception of Risk: An Analysis of Contemporary Worldviews and Cultural Biases. *Journal of Cross-Cultural Psychology*, 22(1), 61–82. <http://doi.org/10.1177/0022022191221006>

Dale, R., & Gore, E. (2010). Process Models and the Development of Trustworthy Digital Repositories. *Information Standards Quarterly*, 22(2), 14. <http://doi.org/10.3789/isqv22n2.2010.02>

Dappert, A. (2009). Report on the Conceptual Aspects of Preservation, Based on Policy and Strategy Models for Libraries, Archives and Data Centres (No. IST-2006-033789). PLANETS-Project. Retrieved from [http://www.planets-project.eu/docs/reports/Planets\\_PP2\\_D3\\_ReportOnPolicyAndStrategyModelsM36\\_Ext.pdf](http://www.planets-project.eu/docs/reports/Planets_PP2_D3_ReportOnPolicyAndStrategyModelsM36_Ext.pdf)

Dappert, A., & Farquhar, A. (2009). Modelling Organizational Preservation Goals to Guide Digital Preservation. *International Journal of Digital Curation*, 4(2), 119–134. <http://doi.org/10.2218/ijdc.v4i2.102>

- Data Seal of Approval Board. (2013a). Data Seal of Approval Guidelines, version 2 (Guidelines). Retrieved from [http://datasealofapproval.org/media/filer\\_public/2013/09/27/guidelines\\_2014-2015.pdf](http://datasealofapproval.org/media/filer_public/2013/09/27/guidelines_2014-2015.pdf)
- Data Seal of Approval Board. (2013b). Data Seal of Approval Regulations (Guidelines). Retrieved from [http://datasealofapproval.org/media/filer\\_public/2013/09/27/dsa-regulations\\_2013.pdf](http://datasealofapproval.org/media/filer_public/2013/09/27/dsa-regulations_2013.pdf)
- Data Seal of Approval. (2014a). About Data Seal of Approval. Retrieved July 19, 2014, from <http://datasealofapproval.org/en/information/about>
- Data Seal of Approval. (2014b). Data Seal of Approval Assessment. Retrieved July 19, 2014, from <http://datasealofapproval.org/en/assessment/>
- Data Seal of Approval. (2015). Data Seal of Approval Community. Retrieved May 14, 2015, from <http://datasealofapproval.org/en/community/>
- Data Seal of Approval. (2016). Community [Nonprofit]. Retrieved May 2, 2016, from <http://datasealofapproval.org/en/community/>
- Day, M. (2008). Toward Distributed Infrastructures for Digital Preservation: The Roles of Collaboration and Trust. *International Journal of Digital Curation*, 3(1), 15–28. <http://doi.org/10.2218/ijdc.v3i1.39>
- De Santis, L., Scannapieco, M., & Catarci, T. (2003). Trusting Data Quality in Cooperative Information Systems. *Lecture Notes in Computer Science*, 2888, 354–369.
- De Vorse, K., & McKinney, P. (2010). Digital Preservation in Capable Hands: Taking Control of Risk Assessment at the National Library of New Zealand. *Information Standards Quarterly*, 22(2), 41–44.
- Denzin, N. K. (1997). *Interpretive Ethnography: Ethnographic Practices for the 21st Century*. Thousand Oaks, CA: Sage Publications.
- Dobratz, S., & Schoger, A. (2007). Trustworthy Digital Long-Term Repositories: The nestor Approach in the Context of International Developments. In L. Kovács, N. Fuhr, & C. Meghini (Eds.), *Research and Advanced Technology for Digital Libraries* (pp. 210–222). Springer Berlin Heidelberg. Retrieved from [http://link.springer.com/chapter/10.1007/978-3-540-74851-9\\_18](http://link.springer.com/chapter/10.1007/978-3-540-74851-9_18)
- Douglas, M., & Wildavsky, A. (1982). *Risk and Culture: An Essay on the Selection of Technological and Environmental Dangers*. Berkeley, CA: University of California Press.
- Dryden, J. (2011). Measuring Trust: Standards for Trusted Digital Repositories. *Journal of Archival Organization*, 9(2), 127–130.
- Federal Depository Library Program. (2014, December 22). GPO Prepares to Become First Federal Agency Named as Trustworthy Digital Repository for Government Information [Government]. Retrieved from <http://www.fdlp.gov/news-and-events/2158-gpo-prepares-to->



become-first-federal-agency-named-as-trustworthy-digital-repository-for-government-information

Fenton, N., Passey, A., & Hems, L. (1999). Trust, the Voluntary Sector and Civil Society. *International Journal of Sociology and Social Policy*, 19(7/8), 21–42. <http://doi.org/10.1108/01443339910788848>

Feynman, R. P. (1988). An Outsider's Inside View of the Challenger Inquiry. *Physics Today*, 41(2), 26. <http://doi.org/10.1063/1.881143>

Fischhoff, B. (1983). *Acceptable Risk*. Cambridge; NY: Cambridge University Press.

Fischhoff, B., Hope, C., & Watson, S. R. (1990). Defining Risk. In T. S. Glickman & M. Gough (Eds.), *Readings in Risk*. Washington, DC: Resources for the Future.

Flynn, J., Slovic, P., & Mertz, C. K. (1994). Gender, Race, and Perception of Environmental Health Risks. *Risk Analysis*, 14(6), 1101–1108. <http://doi.org/10.1111/j.1539-6924.1994.tb00082.x>

Fox, A. J., Gardner, M. J., Lees, F., Green, C., & Andrews, D. (1981). Mortality Statistics and the Assessment of Risk [and Discussion]. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 376(1764), 65–78.

Frank, R. D., & Yakel, E. (2013). Disaster Planning for Digital Repositories. In *Proceedings of the 76th ASIS&T Annual Meeting (Vol. 50)*. Montreal, QC, Canada. Retrieved from <http://asis.org/asist2013/proceedings/submissions/papers/59paper.pdf>

Free, D. (2011). HathiTrust Certified Trustworthy Repository. *College & Research Libraries News*, 72(5), 254.

Gardoni, P., & Murphy, C. (2013). A Scale of Risk. *Risk Analysis*. <http://doi.org/10.1111/risa.12150>

Garrett, J., & Waters, D. J. (1996). *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information (No. 9781887334501 1887334505)* (p. 68). Washington, DC: The Commission on Preservation and Access & Research Libraries Group. Retrieved from <https://www.clir.org/wp-content/uploads/sites/6/pub63watersgarrett.pdf>

Getting Content Into HathiTrust. (2016). [Nonprofit]. Retrieved April 7, 2016, from <https://www.hathitrust.org/ingest>

Giaretta, D., Conrad, M., Garrett, J., Longstreth, T., Lambert, S., Sierman, B., ... Tibbo, H. (2011). *Audit and Certification Process for Digital Repositories*. Presented at the PV2011, Toulouse, France. Retrieved from <http://www.iso16363.org/assets/PV2011Giaretta.pdf>

Giaretta, D., Harmsen, H., & Keitel, C. (2010, August 7). *Memorandum of Understanding to create a European Framework for Audit and Certification of Digital Repositories*. Retrieved from

[http://datasealofapproval.org/media/filer\\_public/2014/08/28/20100709\\_020\\_signedmoutocreatea europeanframeworkforauditandcertificationofdigitalrepositories.pdf](http://datasealofapproval.org/media/filer_public/2014/08/28/20100709_020_signedmoutocreatea europeanframeworkforauditandcertificationofdigitalrepositories.pdf)

Governance. (2016). [Nonprofit]. Retrieved April 7, 2016, from <https://www.hathitrust.org/governance>

Groves, R. M., Fowler, F. J., Couper, M., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. Hoboken, NJ: Wiley.

Gubrium, J. F., & Holstein, J. A. (2001). *Handbook of Interview Research*. Thousand Oaks, CA: SAGE Publications, Inc.

Harmsen, H. (2008). Data Seal of Approval - Assessment and Review of the Quality of Operations for Research Data Repositories (pp. 220–223). Presented at the iPRES 2008: The Fifth International Conference on Preservation of Digital Objects, London, U.K.: The British Library. Retrieved from [http://www.bl.uk/ipres2008/presentations\\_day2/34\\_Harmsen.pdf](http://www.bl.uk/ipres2008/presentations_day2/34_Harmsen.pdf)

Hart, P. E., & Liu, Z. (2003). Trust in the Preservation of Digital Information. *Communications of the ACM*, 46(6), 93–97. <http://doi.org/10.1145/777313.777319>

Hilgartner, S. (1992). The Social Construction of Risk Objects. In J. F. Short, Jr. & L. Clarke (Eds.), *Organizations, Uncertainties, and Risk* (pp. 39–53). Boulder, CO: Westview Press.

Hitchcock, J. L. (2001). Gender Differences in Risk Perception: Broadening the Contexts. *Risk: Health, Safety & Environment*, 12, 179.

Hitchcock, S., Brody, T., Hey, J. M. N., & Carr, L. (2007). Digital Preservation Service Provider Models for Institutional Repositories: Towards Distributed Services. *D-Lib Magazine*, 13(5/6).

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley Pub. Co. Retrieved from <http://mirlyn.lib.umich.edu/Record/000000426>

Houghton, B. (2015). Trustworthiness: Self-assessment of an Institutional Repository against ISO 16363-2012. *D-Lib Magazine*, 21(3/4). <http://doi.org/10.1045/march2015-houghton>

Hughes, R. (2004). Vignette Technique. In M. Lewis-Beck, A. Bryman, & T. Liao, *Encyclopedia of Social Science Research Methods*. Thousand Oaks, CA: SAGE Publications, Inc.

Hutter, B. M. (2005). “Ways of Seeing”: Understandings of Risk in Organizational Settings. In B. M. Hutter & M. Power (Eds.), *Organizational Encounters with Risk* (pp. 67–91). Cambridge: Cambridge University Press.

Hutter, B. M., & Power, M. (2005a). *Organizational Encounters with Risk*. Cambridge: Cambridge University Press. Retrieved from <http://dx.doi.org/10.1017/CBO9780511488580>

Hutter, B. M., & Power, M. (2005b). *Organizational Encounters with Risk: An Introduction*. In B. M. Hutter & M. Power (Eds.), *Organizational Encounters with Risk* (pp. 1–32). Cambridge: Cambridge University Press.

- Innocenti, P., McHugh, A., & Ross, S. (2008). Tackling the Risk Challenge: Drambora (digital Repository Audit Method Based on Risk Assessment). In eChallenges 2008. Stockholm. Retrieved from <http://www.iospress.nl>
- ISO-PTAB. (2011). PTAB Home. Retrieved August 11, 2014, from <http://www.iso16363.org/>
- ITHAKA. (2015a). About Us [Nonprofit]. Retrieved April 7, 2016, from <http://www.portico.org/digital-preservation/about-us>
- ITHAKA. (2015b). Facts & Figures [Nonprofit]. Retrieved April 7, 2016, from <http://www.portico.org/digital-preservation/the-archive-content-access/archive-facts-figures#page-1554>
- ITHAKA. (2015c). Governance [Nonprofit]. Retrieved April 7, 2016, from <http://www.portico.org/digital-preservation/about-us/advisory-committee>
- ITHAKA. (2015d). Perpetual Access [Nonprofit]. Retrieved April 7, 2016, from <http://www.portico.org/digital-preservation/the-archive-content-access/perpetual-access>
- Jantz, R., & Giarlo, M. (2007). Digital Archiving and Preservation: Technologies and Processes for a Trusted Repository. *Journal of Archival Organization*, 4(1-2), 193–213. [http://doi.org/10.1300/J201v04n01\\_10](http://doi.org/10.1300/J201v04n01_10)
- Jasanoff, S. (1986). *Risk Management and Political Culture: A Comparative Study of Science in the Policy Context*. NY: Russell Sage Foundation.
- Jasanoff, S. (1998). The Political Science of Risk Perception. *Reliability Engineering & System Safety*, 59(1), 91–99. [http://doi.org/10.1016/S0951-8320\(97\)00129-4](http://doi.org/10.1016/S0951-8320(97)00129-4)
- Kahneman, D. (2013). *Thinking, Fast and Slow* (1st pbk. ed). NY: Farrar, Straus and Giroux.
- Kaplan, S., & Garrick, B. J. (1981). On The Quantitative Definition of Risk. *Risk Analysis*, 1(1), 11–27. <http://doi.org/10.1111/j.1539-6924.1981.tb01350.x>
- Kasperson, R. E., & Kasperson, J. X. (1996). The Social Amplification and Attenuation of Risk. *Annals of the American Academy of Political and Social Science*, 545, 95–105.
- Keitel, C. (2012). DIN Standard 31644 and Nestor Certification. Fondazione Rinascimento Digitale.
- Keitel, C. (2014, October). Nestor Seal (DIN 31644). Slide Deck presented at the DASISH - Workshop on Trust and Certification. Retrieved from [http://dasish.eu/dasishevents/wstrustcertification/2014\\_10\\_17\\_DASISH\\_trust-ws\\_DIN\\_31644\\_nestorSeal\\_Keitel.pdf](http://dasish.eu/dasishevents/wstrustcertification/2014_10_17_DASISH_trust-ws_DIN_31644_nestorSeal_Keitel.pdf)
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in Digital Information. *Journal of the American Society for Information Science and Technology*, 59(3), 363–374. <http://doi.org/10.1002/asi.20722>

- Kirchhoff, A., Fenton, E., Orphan, S., & Morrissey, S. (2010). Becoming a Certified Trustworthy Digital Repository: The Portico Experience. In *Proceedings of the 7th International Conference on Preservation of Digital Objects* (pp. 87–94). Vienna, Austria. Retrieved from <http://ifs.tuwien.ac.at/dp/ipres2010/papers/Kirchhoff-35.pdf>
- Konheim, C. S. (1988). Risk Communication in the Real World. *Risk Analysis*, 8(3), 367–373. <http://doi.org/10.1111/j.1539-6924.1988.tb00499.x>
- Kramer, R. M. (1999). Trust and Distrust in Organizations: Emerging Perspectives, Enduring Questions. *Annual Review of Psychology*, 50(1), 569–598. <http://doi.org/10.1146/annurev.psych.50.1.569>
- Lachlan, K. A., Burke, J., Spence, P. R., & Griffin, D. (2009). Risk Perceptions, Race, and Hurricane Katrina. *Howard Journal of Communications*, 20(3), 295–309. <http://doi.org/10.1080/10646170903070035>
- Lavoie, B. (2008). The Fifth Blackbird: Some Thoughts on Economically Sustainable Digital Preservation. *D-Lib Magazine*, 14(3/4). <http://doi.org/10.1045/march2008-lavoie>
- Lavoie, B., & Dempsey, L. (2004). Thirteen Ways of Looking at...Digital Preservation. *D-Lib Magazine*, 10(7/8). <http://doi.org/10.1045/july2004-lavoie>
- Lawrence, G. W., Kehoe, W. R., Rieger, O. Y., Walters, W. H., Kenney, A. R., & Council on Library and Information Resources, W., DC. (2000). *Risk Management of Digital Information: A File Format Investigation*. Washington, D.C: Council on Library and Information Resources. Retrieved from: <https://www.clir.org/wp-content/uploads/sites/6/pub93.pdf>.
- Lazorchak, B. (2011, August 23). Digital Preservation, Digital Curation, Digital Stewardship: What's in (Some) Names? Retrieved from <http://blogs.loc.gov/digitalpreservation/2011/08/digital-preservation-digital-curation-digital-stewardship-what%E2%80%99s-in-some-names/>
- Leveson, N., Dulac, N., Marais, K., & Carroll, J. (2009). Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems. *Organization Studies*, 30(2-3), 227–249. <http://doi.org/10.1177/0170840608101478>
- MacNeil, H. (2000). Providing the Grounds for Trust: Developing Conceptual Requirements for the Long-Term Preservation of Authentic Electronic Records. *Archivaria*, 50, 52–78.
- Maniatis, P., Roussopoulos, M., Giuli, T. J., Rosenthal, D. S. H., & Baker, M. (2005). The LOCKSS Peer-to-Peer Digital Preservation System. *ACM Transactions on Computing Systems*, 23(1), 2–50. <http://doi.org/10.1145/1047915.1047917>
- McHugh, A. (2012). A Model for Digital Preservation Repository Risk Relationships. In *World Library Information Congress: 78th IFLA General Conference and Assembly*. Helsinki, Finland. Retrieved from <http://eprints.gla.ac.uk/65420>

- McHugh, A., Ross, S., Innocenti, P., Ruusalepp, R., & Hoffman, H. (2008). Bringing Self-Assessment Home: Repository Profiling and Key Lines of Enquiry within DRAMBORA. *The International Journal of Digital Curation*, 3(2), 130–142. <http://doi.org/doi:10.2218/ijdc.v3i2.64>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks: Sage Publications.
- Minor, D., Sutton, D., Kozbial, A., Westbrook, B., Burek, M., & Smorul, M. (2010). Chronopolis Digital Preservation Network. *International Journal of Digital Curation*, 5(1), 119–133. <http://doi.org/10.2218/ijdc.v5i1.147>
- Murphy, M. (2006). *Sick Building Syndrome and the Problem of Uncertainty: Environmental Politics, Technoscience, and Women Workers*. Durham, NC: Duke University Press.
- Mutula, S. M. (2011). Ethics and Trust in Digital Scholarship. *The Electronic Library*, 29(2), 261–276. <http://doi.org/10.1108/02640471111125212>
- Nelkin, D. (1989). Communicating Technological Risk: The Social Construction of Risk Perception. *Annual Review of Public Health*, 10(1), 95–113. <http://doi.org/10.1146/annurev.pu.10.050189.000523>
- nestor Certification Working Group. (2013). *Explanatory Notes on the nestor Seal for Trustworthy Digital Archives (No. nestor-materials 17)*. Frankfurt am Main: Deutsche Nationalbibliothek. Retrieved from [http://files.d-nb.de/nestor/materialien/nestor\\_mat\\_17\\_eng.pdf](http://files.d-nb.de/nestor/materialien/nestor_mat_17_eng.pdf)
- Nestor Seal for Trustworthy Digital Archives. (2013, June 8). Retrieved February 11, 2014, from [http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel\\_node.html](http://www.langzeitarchivierung.de/Subsites/nestor/EN/nestor-Siegel/siegel_node.html)
- nestor Working Group Trusted Repositories - Certification. (2009). *nestor Criteria: Catalogue of Criteria for Trusted Digital Repositories, Version 2*. Frankfurt am Main: Deutsche Nationalbibliothek. Retrieved from <http://nbn-resolving.de/urn:nbn:de:0008-2010030806>
- Newman, I., & Benz, C. R. (1998). *Qualitative-Quantitative Research Methodology: Exploring the Interactive Continuum*. Southern Illinois University Press.
- Nickel, P. J., & Vaesen, K. (2012). Risk and Trust. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of Risk Theory* (pp. 857–876). Springer Netherlands.
- Nolan, F. P. (2007). *Technology and Trust: Constituency Relationships in the Online Environment*. Retrieved from *Sociological Abstracts*. (61665708; 200723723)
- Ohshima, S. (2010). Risks Associated with Digital Preservation: Media Deterioration, Media Obsolescence and File Format Obsolescence. *Journal of Information Science & Technology Association/Joho No Kagaku to Gijutsu*, 60(2), 54–54.
- Olofsson, A., Zinn, J. O., Griffin, G., Nygren, K. G., Cebulla, A., & Hannah-Moffat, K. (2014). The Mutual Constitution of Risk and Inequalities: Intersectional Risk Theory. *Health, Risk & Society*, 0(0), 1–14. <http://doi.org/10.1080/13698575.2014.942258>

- Ontario Council of University Libraries. (2014a). Directors. Retrieved April 7, 2016, from <http://ocul.on.ca/node/38>
- Ontario Council of University Libraries. (2014b). Preservation. Retrieved April 7, 2016, from <http://ocul.on.ca/node/101>
- Ontario Council of University Libraries. (2014c). Scholars Portal. Retrieved April 7, 2016, from <http://ocul.on.ca/node/135>
- Ontario Council of University Libraries. (2016). Scholars Portal Homepage. Retrieved April 7, 2016, from <http://scholarsportal.info/>
- Otway, H. (1992). Public Wisdom, Expert Fallibility: Toward a Contextual Theory of Risk. In S. Krimsky & D. Golding (Eds.), *Social Theories of Risk* (pp. 215–228). Westport, CT: Praeger Publishers.
- Parthasarathy, S. (2004). Regulating Risk: Defining Genetic Privacy in the United States and Britain. *Science, Technology & Human Values*, 29(3), 332–352. <http://doi.org/10.1177/0162243904264485>
- Parthasarathy, S. (2007). Building Genetic Medicine: Breast Cancer, Technology, and the Comparative Politics of Health Care. Retrieved from <http://site.ebrary.com/id/10173530>
- Pearce, D. W., Russell, S., & Griffiths, R. F. (1981). Risk Assessment: Use and Misuse [and Discussion]. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 376(1764), 181–192.
- Perrow, C. (1999). *Normal Accidents: Living with High-Risk Technologies* (Updated). Princeton University Press.
- Pidgeon, N. (1998). Risk Assessment, Risk Values and the Social Science Programme: Why We Do Need Risk Perception Research. *Reliability Engineering & System Safety*, 59(1), 5–15. [http://doi.org/10.1016/S0951-8320\(97\)00114-2](http://doi.org/10.1016/S0951-8320(97)00114-2)
- Pirson, M., & Malhotra, D. (2011). Foundations of Organizational Trust: What Matters to Different Stakeholders? *Organization Science*, 22(4), 1087–1104. <http://doi.org/10.1287/orsc.1100.0581>
- Prior, L. (2003). Chapter One: Basic Themes: Use, Production and Content. In *Using documents in social research* (pp. 1–29). London, UK: SAGE.
- PTAB Courses. (2015). Retrieved May 14, 2015, from <http://www.iso16363.org/courses>
- PTAB. (2014). Retrieved April 8, 2016, from <http://www.iso16363.org/ptabmembers/>
- Reich, V., & Rosenthal, D. S. H. (2001). LOCKSS: A Permanent Web Publishing and Access System. *D-Lib Magazine*, 7(6). <http://doi.org/10.1045/june2001-reich>

- Renn, O. (1991). Risk Communication and the Social Amplification of Risk. In R. E. Kasperson & P. J. M. Stallen (Eds.), *Communicating Risks to the Public* (pp. 287–324). Springer Netherlands.
- Renn, O. (1999). A Model for an Analytic–Deliberative Process in Risk Management. *Environmental Science & Technology*, 33(18), 3049–3055. <http://doi.org/10.1021/es981283m>
- Renn, O. (2008). White Paper on Risk Governance: Toward an Integrative Framework. In O. Renn & K. D. Walker (Eds.), *Global Risk Governance* (pp. 3–73). Springer Netherlands.
- Renn, O., Burns, W. J., Kasperson, J. X., Kasperson, R. E., & Slovic, P. (1992). The Social Amplification of Risk: Theoretical Foundations and Empirical Applications. *Journal of Social Issues*, 48(4), 137–160. <http://doi.org/10.1111/j.1540-4560.1992.tb01949.x>
- Rijpma, J. A. (1997). Complexity, Tight–Coupling and Reliability: Connecting Normal Accidents Theory and High Reliability Theory. *Journal of Contingencies and Crisis Management*, 5(1), 15–23. <http://doi.org/10.1111/1468-5973.00033>
- RLG-NARA Digital Repository Certification Task Force. (2007). *Trustworthy Repositories Audit & Certification: Criteria and Checklist, Version 1.0*. Retrieved from [http://www.crl.edu/sites/default/files/attachments/pages/trac\\_0.pdf](http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf)
- RLG-OCLC Working Group on Digital Archive Attributes. (2002). *Trusted Digital Repositories: Attributes and Responsibilities*. Mountain View, CA: Research Libraries Group (RLG). Retrieved from <http://www.oclc.org/programs/ourwork/past/trustedrep/repositories.pdf>
- Rosenthal, D. S. H. (2010). Format obsolescence: assessing the threat and the defenses. *Library Hi Tech*, 28(2), 195–210. <http://doi.org/10.1108/07378831011047613>
- Ross, S., & McHugh, A. (2006a). *Preservation Pressure Points: Evaluating Diverse Evidence for Risk Management*. Presented at the iPRES 2006, New York, NY: Digital Curation Centre.
- Ross, S., & McHugh, A. (2006b). The Role of Evidence in Establishing Trust in Repositories. *D-Lib Magazine*, 12(7/8). <http://doi.org/10.1045/july2006-ross>
- Rousseau, D. M., Sitkin, S. B., Burt, R. S., & Camerer, C. (1998). Not so Different After All: A Cross-Discipline View of Trust. *Academy of Management Review*, 23(3), 393–404.
- Routhier Perry, S. (2014). Digitization and Digital Preservation: A Review of the Literature. *SLIS Student Research Journal*, 4(1). Retrieved from <http://scholarworks.sjsu.edu/slissrj/vol4/iss1/4>
- Rowe, W. D. (1977). *An Anatomy of Risk*. NY: Wiley.
- Royal Society (Great Britain), & Study Group on Risk. (1983). *Risk Assessment: Report of a Royal Society Study Group*. London: Royal Society.

Saldaña, J. (2015). *The Coding Manual for Qualitative Researchers* (3rd edition). Thousand Oaks, CA: Sage Publications.

Saldaña, J. (2015). *Thinking Qualitatively: Methods of Mind*. Thousand Oaks, CA: Sage Publications.

Schultz, M., & Gore, E. (2010). The Importance of Trust in Distributed Digital Preservation: A Case Study from the Metaarchive Cooperative. In *Proceedings of the 7th International Conference on Preservation of Digital Objects* (pp. 105–112). Vienna, Austria. Retrieved from <http://ifs.tuwien.ac.at/dp/ipres2010/papers/schultz-39.pdf>

Silver, N. (2012). *The Signal and the Noise: Why So Many Predictions Fail--But Some Don't*. NY: Penguin Press.

Sjöberg, L. (2000). Factors in Risk Perception. *Risk Analysis*, 20(1), 1–12.  
<http://doi.org/10.1111/0272-4332.00001>

Slovic, P. (1987). Perception of Risk. *Science*, 236(4799), 280–285.  
<http://doi.org/10.1126/science.3563507>

Smith Rumsey, A. (2016). *When We Are No More: How Digital Memory Is Shaping Our Future*.

Starr, C. (1969). Social Benefit versus Technological Risk. *Science*, 165(3899), 1232–1238.  
<http://doi.org/10.1126/science.165.3899.1232>

Starr, C. (2003). The Precautionary Principle Versus Risk Analysis. *Risk Analysis*, 23(1), 1–3.  
<http://doi.org/10.1111/1539-6924.00285>

Statistics Information. (2016). [Nonprofit]. Retrieved April 7, 2016, from [https://www.hathitrust.org/statistics\\_info](https://www.hathitrust.org/statistics_info)

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, 13(1). Retrieved from <http://aisel.aisnet.org/cais/vol13/iss1/24>

Strodl, S., Becker, C., Neumayer, R., & Rauber, A. (2007). How to Choose a Digital Preservation Strategy: Evaluating a Preservation Planning Procedure. In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 29–38). New York, NY: ACM.  
<http://doi.org/10.1145/1255175.1255181>.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco, CA: Jossey-Bass Publishers.

The Data Seal of Approval Board. (2011). *Implementation of the Data Seal of Approval: Inter-University Consortium for Political and Social Research* (p. 20). Data Seal of Approval. Retrieved from <http://assessment.datasealofapproval.org/seals/>



Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The Psychology of Survey Response*. Cambridge, UK; NY: Cambridge University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.

Tyler, T. R., & Kramer, R. M. (1995). *Trust in Organizations: Frontiers of Theory and Research*. SAGE.

U-M Office of Budget and Planning. (2016). *Budget Detail (Grey Book) (Budget)*. Ann Arbor, MI: University of Michigan. Retrieved from <http://obp.umich.edu/root/budget/budget-detail/>

van Est, R., Walhout, B., & Brom, F. (2012). Risk and Technology Assessment. In S. Roeser, R. Hillerbrand, P. Sandin, & M. Peterson (Eds.), *Handbook of Risk Theory* (pp. 1067–1091). Springer Netherlands. Retrieved from [http://link.springer.com.proxy.lib.umich.edu/referenceworkentry/10.1007/978-94-007-1433-5\\_43](http://link.springer.com.proxy.lib.umich.edu/referenceworkentry/10.1007/978-94-007-1433-5_43)

Vaughan, D. (1996). *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. Chicago, IL: University of Chicago Press.

Vaughan, D. (2005). Organizational Rituals of Risk and Error. In B. M. Hutter & M. Power (Eds.), *Organizational Encounters with Risk* (pp. 33–66). Cambridge: Cambridge University Press.

Vaughan, E., & Seifert, M. (1992). Variability in the Framing of Risk Issues. *Journal of Social Issues*, 48(4), 119–135. <http://doi.org/10.1111/j.1540-4560.1992.tb01948.x>

Vermaaten, S., Lavoie, B., & Caplan, P. (2012). Identifying Threats to Successful Digital Preservation: the SPOT Model for Risk Assessment. *D-Lib Magazine*, 18(9/10). <http://doi.org/10.1045/september2012-vermaaten>

Viscusi, W. K. (1985). A Bayesian perspective on biases in risk perception. *Economics Letters*, 17(1-2), 59–62. [http://doi.org/10.1016/0165-1765\(85\)90127-2](http://doi.org/10.1016/0165-1765(85)90127-2)

Wildavsky, A., & Dake, K. (1990). Theories of Risk Perception: Who Fears What and Why? *Daedalus*, 119(4), 41–60.

Wilkinson, I. (2001). Social Theories of Risk Perception: At Once Indispensable and Insufficient. *Current Sociology*, 49(1), 1–22. <http://doi.org/10.1177/0011392101049001002>

Wohlers, A. E. (2010). Regulating Genetically Modified Food: Policy Trajectories, Political Culture, and Risk Perceptions in the U.S., Canada, and EU. *Politics and the Life Sciences*, 29(2), 17–39. [http://doi.org/10.2990/29\\_2\\_17](http://doi.org/10.2990/29_2_17)

Wright, R., Miller, A., & Addis, M. (2009). The Significance of Storage in the “Cost of Risk” of Digital Preservation. *International Journal of Digital Curation*, 4(3), 104–122. <http://doi.org/10.2218/ijdc.v4i3.125>

Wynne, B. (1992). Misunderstood Misunderstanding: Social Identities and Public Uptake of Science. *Public Understanding of Science*, 1(3), 281–304. <http://doi.org/10.1088/0963-6625/1/3/004>

Yakel, E. (2007). Digital Curation. *OCLC Systems & Services: International Digital Library Perspectives*, 23(4), 335–340. <http://doi.org/10.1108/10650750710831466>

Yearley, S. (2000). Making Systematic Sense of Public Discontents with Expert Knowledge: Two Analytical Approaches and a Case Study. *Public Understanding of Science*, 9(2), 105–122. <http://doi.org/10.1088/0963-6625/9/2/302>

Yin, R. K. (2003). *Case Study Research: Design and Methods*. Thousand Oaks, CA: Sage Publications.

Yoon, A. (2014). End Users' Trust in Data Repositories: Definition and Influences on Trust Development. *Archival Science*, 14(1), 17–34. <http://doi.org/10.1007/s10502-013-9207-8>